ONDOKUZ MAYIS ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ
ANABİLİM DALI

A CORPUS BASED STUDY ON VOCABULARY OF AN A2 LEVEL ENGLISH COURSEBOOK

Yüksek Lisans Tezi

Doğan BAYDAL

Danışman:
Doç. Dr. Nalan KIZILTAN

Samsun, 2016

ONDOKUZ MAYIS ÜNİVERSİTESİ
EĞİTİM BİLİMLERİ ENSTİTÜSÜ
YABANCI DİLLER EĞİTİMİ
ANABİLİM DALI

A CORPUS BASED STUDY ON VOCABULARY OF AN A2 LEVEL ENGLISH COURSEBOOK

Yüksek Lisans Tezi

Doğan BAYDAL

Danışman:

Doç. Dr. Nalan KIZILTAN

Samsun, 2016

BİLİMSEL ETİK BİLDİRİMİ

Hazırladığım Yüksek Lisans çalışmasının bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu, yazımda enstitü yazım kılavuzuna uygun davranıldığını taahhüt ederim.

05 / 09 / 2016

Doğan BAYDAL

KABUL VE ONAY

Doğan BAYDAL tarafından hazırlanan *A Corpus Based Study On Vocabulary Of An A2 Level English Coursebook* başlıklı bu çalışma, 05.09.2016 tarihinde yapılan savunma sınavı sonucunda oy birliğiyle başarılı bulunarak jürimiz tarafından Yüksek Lisans Tezi olarak kabul edilmiştir.

İMZA

Başkan:	Doç. Dr. Arif SARIÇOBAN	-----------------------------
Üye:	Doç. Dr. Nalan KIZILTAN	-----------------------------
Üye:	Yrd. Doç. Dr. Dilek ÇAKICI	-----------------------------

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

… / … /2016

Prof. Dr. Önder KABADAYI

Enstitü Müdürü

ÖZET

A2 SEVİYESİNDEKİ İNGİLİZCE DERS KİTABI ÜZERİNDE

DERLEM TEMELLİ ÇALIŞMA

Doğan BAYDAL

Ondokuz Mayıs Üniversitesi, Eğitim Bilimleri Enstitüsü

Yabancı Diller Eğitimi Ana Bilim Dalı, Yüksek Lisans, Eylül / 2016

Danışman: Doç. Dr. Nalan KIZILTAN

Bu çalışmanın amacı, A2 düzeyindeki İngilizce ders kitabındaki sözcüklerin kullanımını, en sık kullanılan sözcükler, bu sözcüklerin Britanya Ulusal Derlemine uygunluğu, Avrupa Dilleri Ortak Çerçeve Metni A2 düzeyine uygunluğu ve 8. Sınıf öğrencilerinin ilgileriyle örtüşüp örtüşmemesi açısından, incelemektir. Bu çalışma ayrıca ders kitabı yazarları ve öğretmenlerin yabancı dil öğretiminde sözcük seçimine dikkatlerini çekmeyi amaçlamaktadır. Ders kitabındaki tüm sözcükler iki farklı derlem yazılım programı tarafından çözümlenmiştir. Ders kitabındaki sözcüklerin, 8. Sınıf öğrencilerinin ilgi alanları ile örtüşüp örtüşmediğini incelemek için ise Türkiye'nin yedi faklı bölgesinden 527 tane 8. Sınıf öğrencisine ilgi alanlarını belirtmelerini isteyen sormaca uygulanmıştır. Bu çalışmanın sonuçlarına göre, kitaptaki sözcükler İngilizcede en çok kullanılan 1000 sözcük ile örtüşmektedir (76%), fakat İngilizcede en çok kullanılan ikinci 1000 sözcüğün ders kitabındaki oranı çok azdır (9.10%). İngilizcede sık kullanılmayan sözcüklerin ders kitabında sık kullanıldığı görülmüştür (12.09%); ancak akademik sözcükler fazla yer almamaktadır (3.04%). Ders kitabındaki sözcüklerin ünitelere göre dağılımında ise dört ulamın (İngilizcede en sık kullanılan 1000 kelime, en sık kullanılan ikinci 1000 kelime, sık kullanılmayan sözcükler ve akademik sözcükler) orantısız bir şekilde dağıldığı görülmüştür. Ayrıca, ders kitabındaki sözcükler; sözcük, sözcük türleri ve anlam alanları açısından Britanya Ulusal Derlemi ile uyuşmamaktadır ve öğrencilerin ilgi odağına yönelik hazırlanmamıştır.

Anahtar Sözcükler: Derlem, sözcük sıklığı, ders kitabı inceleme, Avrupa Dilleri Ortak Çerçeve Metni, öğrenci ilgi alanları

ABSTRACT

A CORPUS BASED STUDY ON VOCABULARY

OF AN A2 LEVEL ENGLISH COURSEBOOK

Doğan BAYDAL

Ondokuz Mayıs University, Institute of Educational Sciences

Foreign Language Education Department, M.A., September / 2016

Supervisor: Assoc. Prof. Dr. Nalan KIZILTAN

The purpose of that study is to analyse the use of vocabulary in an A2 level EFL coursebook in terms of the use of high frequent words, selection of the words in accordance with the British National Corpus, appropriateness of the words to A2 Level of the CEFR and the correspondence of the vocabulary in the coursebook to the interests of the A2 level English Language Learners. It also aims at taking attention of course book writers and teachers in selecting vocabulary in foreign language teaching. The vocabulary profile of the whole words in the coursebook has been conducted by two corpus software programmes. To analyse the correspondence of the vocabulary in the coursebook to the interests of the A2 level English Language Learners, 527 8th grade English Language Learners from seven regions of Turkey have been given questionnaires to state their interests. The results have shown that the words in the coursebook correspond with the 1000 most frequent words in General English (76%), however, the percentage of the second 1000 most frequent words is very low (9.10%). The percentage of the Off-List words is unusually high (12.09%), nevertheless, the percentage of academic words is very low (3.04%). Moreover, when the vocabulary coverage of each unit is analysed separately, it is seen that the percentages of the four categories (1000 most frequent words, second 1000 most frequent words, Off-List Words and Academic Words) do not seem balanced. Moreover, the words in the coursebook do not correspond with the British National Corpus in terms of word, parts of speech and semantic field levels and the words in the coursebook do not match with the interests of the 8th grade learners, either.

Key Words: Corpus, word frequency, coursebook analysis, CEFR, students' interests.

ACKNOWLEDGEMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF GRAPHICS

LIST OF ABBREVIATIONS

BNC   **:** British National Corpus

CEFR   **:** Common European Framework of Reference for Languages

CIC   **:** Cambridge International Corpus

CLT   **:** Communicative Language Teaching

COBUILD   **:** Collins Birmingham University International Language Database

CoE   **:** Council of Europe

EFL   **:** English as a Foreign Language

ELT   **:** English Language Teaching

GSL   **:** General Service List

NGSL   : New General Service List

CHAPTER ONE

INTRODUCTION

The value of vocabulary in language teaching and learning is very important. Laufer (1986) agrees on this idea by stating that if the learner does not have the knowledge of the vocabulary, learning a language cannot be achieved. Laufer (1986) further adds that vocabulary has not been given importance by the practitioners of English Language Teaching (ELT) up to now. Along with the growth of Communicative Language Teaching (CLT), accuracy on grammatical structures lost its importance and communicative competency was favoured, which prioritise the importance of vocabulary.

Lexical syllabus has come into existence as the natural outcome of the Communicative Language Teaching. According to lexical syllabus, the words with high level of frequency should be taught first because they occur with a high percentage in spoken and written English (Thurnbury, 2002). As well as lexical syllabus, the field of corpus linguistics has become popular because of the high frequent words which are measured by corpus software programmes. Word frequency is an important concept of corpus linguistics and it shows how many times a word or word phrase takes place in spoken and written language. Word frequency is a prerequisite of Communicative Language Teaching; it is because if the learner knows the most frequent words in the target language, s/he will be able to understand the meaning of the written and spoken language highly, which means that with a small number of words, fluency in speaking can be achieved. Furthermore, teachers and coursebook writers can get benefit from word frequency in that the word frequency helps them to decide which words to be selected in foreign language teaching.

Nowadays, many coursebooks focus attention on the importance of word frequency. A coursebook *Cutting Edge Intermediate* (2005) emphasizes the importance of vocabulary by focusing on high frequent useful words and phrases.

Then the most frequent words can be great helpers for an English as a Foreign Language (EFL) teacher in order to teach learners the required vocabulary.

1.1. The Problem Statement

English lessons must be planned so as to help the learners to use their communicative competence, which requires appropriate vocabulary choice in their interaction and self-presentation. Thus, teachers and coursebook writers have a heavy burden on their shoulders in choosing the right vocabulary in foreign language teaching.

Word frequency may be the most useful way in choosing the right vocabulary, which requires a research according to the needs of the learners at all levels from different regions. The knowledge of the most frequent words shows that the learner has a large number of vocabulary to understand the written and spoken language. Nation (2006) agrees on that idea by stating that the knowledge of the most common 1000 words means that the learner will be able to understand the 78% to 81% of the written texts and 85% of the spoken texts. If the percentage of the proper nouns is added to the most common 1000 words, then it will be much easier for the learners to comprehend the language samples.

There are low frequent words as well as high frequent words. The low frequent words exist in a text on rare occasions. According to Nation (2000), low frequent words may stand for unusual expression, they may have the similar meanings with high frequent words, they may be out of date, they may belong to a specific language or they may be proper nouns or foreign words.

Teaching low frequent words or working with them in the classroom settings is not meaningful. Therefore, both teachers and coursebook writers need to keep this fact in mind. Coursebook writers need to consult corpora and wordlists while preparing language teaching books and teachers should get help from high frequent words while preparing lesson plans and course materials.

1.2. Purpose of the Study

This study aims at discussing the use of high frequent words in an A2 level EFL coursebook and taking attention of course book writers and teachers to the place of

vocabulary lists based on frequency in language teaching. It also aims at discussing whether the coursebook takes the interest of the language learners into account in choosing appropriate vocabulary according to their language levels.

1.3. Research Questions

This study has been based on the following research questions:

1. What are the most frequent words in coursebook *Upturn in English 8*?

2. How well are K1 Words (most frequent 1000 words), K2 Words (most frequent second 1000 words), academic words and Off-list words used in the coursebook *Upturn in English 8*?

3. Are there any significant differences between the corpus in the coursebook and the British National Corpus?

4. Is the corpus in the coursebook *Upturn in English 8* appropriate for A2 level according to Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)?

5. Do the most frequent words in the coursebook correspond to the interests of A2 level English Language Learners?

1.4. Scope of the Study

The research aims to investigate A2 level coursebook *Upturn in English 8*, which is used compulsorily in all Turkish public secondary schools and is claimed to provide four skills such as reading, listening, speaking and writing in A2 level of the Common European Framework of Reference (CEFR). The book has been accepted to be used for 5 years beginning from 2015-2016 Academic Year. The research also consists of 527 eighth graders who have been learning English since they were 10, and who attend the public schools in different regions of Turkey.

1.5. Limitations of the Study

The study has been limited to a course book, which provides skills referred only in A2 level of The Common European Framework of Reference (CEFR). The Common European Framework divides learners into six levels: A1, A2, B1, B2, C1, C2. The study also has been limited to British National Corpus (BNC) Spoken data in order to get some results from the corpus software programmes Compleat Lexical Tutor (Cobb, 2015) and Vmatrix (Rayson, 2009).

1.6. Definitions of the Terms

A2 Level: It refers to the "basic user of English who can understand sentences and frequently used expressions related to areas of most immediate relevance, can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters, can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate" (Council of Europe, 2001:24).

Common European Framework of Reference (CEFR): "It provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe" (Council of Europe, 2001:1).

Communicative Language Teaching (CLT): "It is an approach to language teaching, which aims broadly to apply the theoretical perspective of the Communicative Approach by making communicative competence the goal of language teaching and by acknowledging the interdependence of language and communication" (Larsen-Freeman, 2000:121).

Corpus (Plural Corpora): "It is a collection of texts, written or spoken, which is stored on a computer" (O'Keeffe, McCarthy, & Carter, 2007: 1).

Course Book: "It is a book that learners use regularly in the lessons and is seen as a resource in achieving aims and objectives that have already been set in terms of learner needs" (Cunningsworth, 1995:7).

Frequency: "It is how often the word occurs in normal use of the language" (Nation, 1997).

Lemma: Francis and Kucera (1982: 1) define it as a "set of lexical forms having the same stem and belonging to the same major word class, differing only in inflection and/or spelling. For example, the verb lemma WALK consists of the words walk, walked, walking and walks" (as cited in Baker, Hardie, & McEnery, 2006:104).

Lexis: All the words in a language.

Token: "A single linguistic unit, most often a word, although depending on the encoding system being used, a single word can be split into more than one token, for example; he's (he +'s)" (Baker, Hardie, & McEnery, 2006:159).

Type: "While the number of tokens in a corpus refers to the total number of words, the number of types refers to the total number of unique words. For example, the word "ship" may occur 177 times in a corpus, but it only counts as one type of word. Types are used in calculating the type/token ratio (a measure of lexical repetition) of a text or corpus" (Baker, Hardie, & McEnery, 2006:162).

Word Family: "It consists of a base word and all its derived and inflected forms that can be understood by a learner without having to learn each from separately" (Bauer & Nation, 1993:253).

Word List: "A list of all of the words that appear in a text or corpus, often useful for dictionary creation" (Baker, Hardie, & McEnery, 2006:169).

CHAPTER TWO

REVIEW OF LITERATURE

2.1. Corpus Studies

The compilation of written and spoken texts is called corpus. The data of the corpus are held in computer (O'Keeffe, McCarthy & Carter, 2007). Currently, the data of the corpus are uploaded on a website and it becomes available to be observed by the researchers. However, when the computers were not that common as in today, the experts of the corpus linguistics used to observe the written materials and it was so hard to collect, organize and save the data for them. Nevertheless, the development of the technology made the computers to be available much more easily and as a result of this, different software programmes on corpus linguistics area were prepared by the specialists of linguistics. With the advent of these corpus software programmes, collecting, organising, calculating and saving the language data have become simple and as well as experts of linguistics, all the people (teachers, students, etc.) have had the chance of accessing the data of any corpus.

It is important to state that a compilation of any language data cannot be considered as corpus because there should be some requirements that have to be fulfilled while creating a corpus. First, the corpus should be planned and the purpose of it should be for linguistics. Second, the corpus should be representative of the data to be analysed.

First of all, as Hunston (2002) states, planning, organisation and linguistic purpose are the core elements of the corpus. These features differ the corpus from library; it is because the books and the data in the libraries are kept only to be read. Hunston (2002) further adds that it is important that the data of the corpus can be used for quantitative and qualitative analyses. For example, a researcher can compose a corpus which is taught in a secondary school. From this corpus, anyone can learn the total number of the most frequent words, which words collocate with each other mostly and so forth. This is called quantitative analysis. The researcher can also compare the corpus of a

coursebook with the reference corpus (e.g. British National Corpus) to make judgements about it. This observation is called qualitative analysis.

Representativeness is another requirement a corpus should have because the linguists judge the usefulness of a corpus by looking at how representative it is. As O'Keeffe, McCarthy and Carter (2007) state, if you want to create a classroom corpus, the corpus should represent all the variables related to classroom such as: Which type of school is it? What qualifications does the teacher have? (experience, education, gender and so forth) What are the genders and ages of the students? What is the number of the students in the classroom? What are the backgrounds of the students? The variables change according to the purpose of the corpus.

2.2. Corpus Techniques

There are two main techniques in corpus studies. These are frequency and concordance. Corpus software programmes, such as Compleat Lexical Tutor (Cobb, 2015) and Vmatrix (Rayson, 2009) make it possible to analyse the frequency and concordance data of any written and spoken texts.

2.2.1. Frequency

Frequency is the most common corpus technique used by the corpus linguistics specialists. The corpus software programmes count the words or the word phrases in any written or spoken text and a total calculation is reached. The method of calculating the words changes according to whether the text is in written form or in spoken form. If the texts are in the written form, the texts are either scanned or typed and then saved to a computer. However, if the texts are in the spoken form, the process gets hard because it is more time consuming when compared to written texts. For example, if a teacher wants to get the most frequent words that his/her students produce in speaking lessons, first of all, s/he needs the record the conversations of the students and then transcribe them. It is really a hard process since the conversation is composed of many other variables apart from words, such as turn-taking, interruptions, noises coming from outside and so forth. Whether it is in written or spoken form, as a result of these processes, one can get the frequency of words to study on.

O'Keeffe, McCarthy and Carter (2007) state that through word frequency lists, it is possible to observe all the words in order of frequency. Frequency lists are effective in identifying probable differences between the corpora. Below are the most frequent ten words of different corpora.

Table 1: Comparison of word frequencies for the ten most frequent words across five different datasets (O'Keeffe et all, 2007:11)

	1 Shop (LCIE)	2 Friends (LCIE)	3 Academic LIBEL	4 Australian Corpus of English	5 CIC newspaper & magazine sub-corpus
Rank order	spoken	spoken	spoken	written	written
1	you	I	the	the	the
2	of	and	and	of	to
3	is	the	of	and	of
4	thanks	to	you	to	a
5	it	was	to	a	and
6	I	you	a	in	in
7	please	it	that	is	is
8	the	like	in	for	for
9	yeah	that	it	that	it

1 Service encounters: a sub-corpus of the Limerick Corpus of Irish English (LCIE) consisting of shop encounters (8,500 words)
2 Friends chatting: a sub-corpus of LCIE, consisting of female friends chatting (40,000 words)
3 Academic English: The Limerick-Belfast Corpus of Academic Spoken English (LIBEL CASE, one million words of academic English[3])
4 Australian casual conversation: the Macquarie Corpus of English (ACE) (one million words of written Australian English)
5 Written British and American English: The Cambridge International Corpus based on a 100,000 word sample of newspaper and magazines from McCarthy (1998: 122–123).

By looking at the table above, one can easily see the difference between the spoken and the written corpora. While the interaction words, such as *yeah*, *please*, *thanks* are among the most frequent ten words in spoken corpora, they are not among the most frequent ten words in academic spoken corpus and written corpora. On the contrary, the articles *the*, *a* and the preposition *of* are among the most frequent words in the written corpora and they are not among the most frequent ones in the spoken corpora.

As well as comparing different corpora, the frequency of the words can be studied within corpus with different texts. Table 2 shows the frequencies of the words *must*, *have to*, *surprisingly* and *incredibly* by comparing the books published in England, *The Times* newspaper and spoken English.

Table 2: Frequencies of must, have to, incredibly and surprisingly across corpora (per million words) (Hunston, 2002:6)

	Books	*The Times*	Spoken
must	683	460	363
have to	419	371	802
Total	1102	831	1165
incredibly	8	10	15
surprisingly	25	29	4
Total	33	39	19

It can be figured out from Table 2 that the sizes of the three corpora are not same. Thus, evaluating these three corpora in terms of per million scale gives more accurate results to make comments about the frequency of the words. As it is seen in Table 2, spoken corpus differs from the corpora of the books and *The Times* newspaper in using the auxiliary verbs which indicates necessity. While *have to* is used more than *must* in spoken corpus, in the corpora of the books and *The Times* newspaper, *must* is used more than *have to*. Moreover, while the adverb *incredibly* is used more than *surprisingly* in spoken corpus, it is not preferred as much as *surprisingly* in the corpora of the books and *The Times* newspaper. As it can be seen, many results can be reached and many judgements can be made by analysing the frequency lists of the words.

2.2.2. Concordance

Concordance presents all of the occurrences of a specific word in the context it belongs to. The searched word is usually presented in the middle and there are a few words on the left and on the right of it (Bauer & Nation, 1993). The searched item may be only one word or strings of words. The searched word is usually highlighted in bold to make it realized easily. Figure 1 below shows a sample concordance of the word *witnessed*.

1. y told Tom Jones that he had never before	**witnessed**	a Cabinet scene like it." Al
2. the early decades of the twentieth century	**witnessed**	an increase in the power o
3. uld be drawn up carefully and signed and	**witnessed**	in a particular way. If you
4. The first attitude has been	**witnessed**	in the 1930s and during o
5. nk had recovered from the breakdown we	**witnessed**	in late 1986 and, despite th
6. fought essentially on national issues and it	**witnessed**	the return not only of a re
7. The last year of Ayliffe's Presidency	**witnessed**	the fulfilment of one of th
8. eneration after the coming of Cyrus which	**witnessed**	the most brilliant speculat
9. dirt, gloom and misery as I never before	**witnessed**	". Queen Victoria had the
10. ood that this small Year Niner has been	**witnessed**	"to and moves on to his ne

Figure 1: Sample concordance of witnessed (Bauer & Nation, 1993:43)

The concordance lines give chances to make observations about the language data. For example, it is clear to judge from Figure 1 that the word *witnessed* goes before an article or the preposition. Figure 1 also shows the different meanings of *witnessed*, such as seeing something to happen, experiencing the important events or changes and so forth.

The presentation of the searched item in the concordance may be different. The most preferred presentation is the alphabetical presentation. In the alphabetical presentation, the words on the right of the searched word is ordered alphabetically as in Figure 1 (Sinclair, 1991) or the words on the left of the searched word is ordered alphabetically as in Figure 2 below.

1. whole process the activity	**is**	obvious enough the nervous
2. communication where the activity	**is**	halted in time if
3. nervous activity of authors	**is**	legendary and the silent
4. mankind's only remaining boast	**is**	that we thought of
5. radio listener his brain	**is**	highly active if he
6. of activity and communication	**is**	only one of them
7. highly active if he	**is**	taking anything in there
8. small sub-section although it	**is**	very important attempts to
9. of it first it	**is**	certainly an intricate and
10. human communication through language **is**		only a small sub-section

Figure 2: Sample concordance of is (Sinclair, 1991)

The data in the concordance can be studied in terms of two headings: collocation and keyword analysis.

According to Hunston (2002), some words tend to occur more often with each other and this concept is called collocation. For example, *blackboard* occurs with the word

teacher more often than with the word *doctor*. It is easy to judge that *blackboard* collocates with *teacher* usually because they belong to the same context and there is a connection between these words. However, it is not that easy to judge the collocational patterns between the words. Thus, a concordance programme is necessary for statistical evaluations in collocations of the words. Below is a sample concordance of the *way*.

```
ether in northern Ireland is no different in a way then em what they were desperately
    you see it? Some of you anyhow? Now in a way 'What Dreams may come' it's not
    subject to study in college in fact it's a way of life and you find this right
    and how could he present things in such a way that he would persuade people.
ul and the purpose of life is to live in such a way that when you die your soul is
t he was obviously he obviously lived a certain way of live and they wanted to know
  lem that they had to deal with in a different way they couldn't deal with it by
asically in football stadium that's the easiest way to describe it. There is a large
sking for you ok I find this the most effective way.  Ok now today em you have as well
speculative because there is no evidence either way.  You can't have evidence about
 e theologian starts from the top and works his way down.  The theologian will have
 rts from the ground so it speaks and works its way up.  The theologian starts from
```

Figure 3: Concordance lines for way from LCIE (O'Keeffe et all, 2007:9)

It is important to state that the searched word in the concordance must be limited to one part of speech because different parts of speech of the same word result in many different meanings and it leads the researcher to reach incorrect data.

It is clearly understood from Figure 3 that the word *way* generally occurs with the preposition *in* and pronoun *a* and they form prepositional phrases, such as *a way*, *in a way*, *such a way*. Moreover, the word way forms adjective phrases, such as *certain way*, *different way*, *easiest way* and *effective way*. This figure shows that by observing the concordance lines, anyone can get a clear idea about the collocations of a word.

Collocational data can be very useful for the researchers (Sinclair, 1991; Shei and Pain, 2000; Hunston, 2002; Deveci, 2004; Sheehan, 2004; O'Keeffe, McCarthy and Carter, 2007; Altıkulaçoğlu, 2010; Çelik, 2011; Macgregor, 2014; Özbay and Kayaoğlu, 2015). First of all, it becomes very easy to find the searched word by means of concordance lines. Otherwise, the researcher has to look through large corpus to find the searched word. Second, collocational data help the researchers to know the different meanings of a particular word. By analysing the colocations of a word, it is easy to guess the real or the metaphoric meanings of a searched item. Third,

collocational lines make it possible to analyse the phraseology of a word. For example, by looking at Figure 3, it is seen that *such a way*, *in a way*, *way up*, *way down* are the most used phraseologies of the word *way*. Finally, by means of collocations, we can learn which different semantic areas a word belongs to. Orpin (1997) exemplifies this as follows. By analysing the collocations of the words "bribe" and "bribery", we come across with different semantic areas. These are:

Words related with negative meaning (*scandal*, *corruption*)

Words related with finance (*dollar*, *money*, *tax*)

Words related with sports (*referee*, *players*)

Words related with law (*charge*, *investigation*)

These different semantic areas of the words "bribe" and "bribery" are useful in understanding both the meanings and the cultural reflections of the words.

Apart from collocation, there is another method in concordance, that is, keyword analysis. Bauer and Nation (1993) state that keywords occur unusually more frequently than it should be when compared to a reference corpus. It may because of that fact that proper nouns are common in the corpus or there are a lot of words belonging to a specific stylistic category (Berber Sardinha, 1995). The aim of the keyword analysis is to compare the texts and corpora between each other. For example, a researcher may compare the small corpus that s/he created with the large corpus like British National Corpus or Cambridge International corpus to see if there are unusually frequent words in his/her corpus, and if so, s/he may investigate the reasons behind it.

2.3. Types of Corpora

The purpose, the degree of the representativeness, the structure and the formation of the corpus determines the type of it. Below are the most used corpus types.

2.3.1. Specialised Corpus

Specialised corpus is developed for a particular purpose (Kennedy, 1998; Hunston, 2002; O'Keeffe, McCarthy and Carter, 2007). Journal articles, essays written by the

students, newspaper news or history books may be examples of specialised corpus. Specialised corpus may even be more special like classroom language or speaking lessons in the schools. There is no end to the level of specialisation; however, there should be some restrictions on how to select the texts. For example, the texts for the specialised corpus may be restricted to a time period, social context or to a group of people. Dialect corpus, regional corpus and learner corpus are types of specialised corpus (Kennedy, 1998).

2.3.2. Dialect Corpus and Regional Corpus

Dialect corpus is a specific spoken corpus, which is formed for the aim of investigating the regional variations in the language. By means of dialect corpus, speakers may be classified according to their dialects and different types of pronunciations may be distinguished. Regional corpus, on the other hand, is a corpus of language from a specific region. Regional corpus is formed in order to carry out studies of regional or sociolinguistic differences (Baker et al., 2006).

2.3.3. Learner Corpus

The texts of the learners, whether spoken or written, form the learner corpus. The aim here is to compare the texts of the students (e.g. essays written by the students) with each other or with a large reference corpus. For example, the texts of the learners of English as a foreign language may be compared to the texts of the native speakers of English. The results of this comparison are very important in language teaching because the teacher has the chance of having real linguistic evidence (Callies, 2013; Hana et al, 2014; Callies and Götz, 2015). There are many learner corpora around the world, however the most widely known is the International Corpus of Learner English (ICLE). It is composed of the texts written by the learners of English from many different parts of the world such as Japan, Turkey, Russia, France and so forth.

2.3.4. General (Reference) Corpus

General corpus consists of texts from many different sources (Bauer & Nation, 1993). General corpus may be composed of written or spoken data, or both, may belong to one-time period or more, may belong to one country or many countries. General corpus

is much larger than a specialised corpus and stands as a reference. Below are some of the most widely known and used general corpora.

2.3.4.1. British National Corpus (BNC)

British National Corpus has 100 million words derived from written and spoken British English. 90% of the BNC have written data while 10% of it have spoken data. Crowdy (1993) explains the process the collection of the spoken and written data of the BNC as follows. The written data of the BNC is composed of books, periodicals booklets, letters, articles, essays and so forth. The spoken data of the BNC, on the other hand, are composed of 10 million spoken examples derived from the conversations in the streets, classrooms, sports organizations, meetings, phone talking, television programmes and so forth. Apart from these, the conversations of 124 volunteers were recorded on daily basis and then transcribed. It is seen that especially the language data of the spoken corpus of the BNC are varied and it have high representativeness, thus stand as the basis among other general corpora.

2.3.4.2. Brown Corpus

Brown Corpus was created by Brown University and stands as the first computer-driven corpus for researchers of linguistics. Brown Corpus has nearly one million words derived from written English in the United States in 1961. Brown Corpus is composed of 500 samples and each sample has nearly 2000 words. These 500 samples contain 15 different text classes (Kennedy; 1998; O'Keeffe et all, 2007).

2.3.4.3. Cambridge International Corpus (CIC)

Cambridge International Corpus has one billion words from British, American and Learner English. British English data have 17 million spoken, 20 million written, 30 million business and 1 million spoken business words. American English data have 200 million written, 22 million spoken, 7 million written academic and 30 million written business words. Learner English data have 19 million written and 12 million error-coded written words. As it is seen, CIC is rich in different kinds of texts.

2.3.5. Comparable Corpus

Comparable corpus is used to judge about the similarities and differences of two or more corpora in different languages or in the same language with different dialects such as Canadian English, Australian English and so forth. Comparable corpora are useful in translation analysis.

2.3.6. Parallel Corpus

Parallel corpus is used to make comparisons between the texts from two different languages. For example, the texts of the international associations are prepared in one language and generally translated into other languages like European Union regulations. In this context, parallel corpora are used to analyse the comparison of the two language (Kennedy; 1998).

2.4. Corpus and Foreign Language Teaching

Throughout the years, linguistic analysis was the core purpose of the corpus studies. However, with the first half of the 1980s, a lot of research have been made about corpus and foreign language teaching and learning. John Sinclair, who is a professor in English department, made the first important study in this field. He created the COBUILD Project (Sinclair, 1987) which was about running the corpus data in computers and working with its implementations in language teaching. Römer (2010) states that the aim of the COBUILD Project was to present useful dictionaries and genuine teaching materials to the learners of English. The COBUILD Project provides language samples which are not invented but authentic. Students face real conversation data and written texts of native speakers of English. Römer's (2010) statements make it clear that COBUILD Project is composed of language samples which are very crucial in language teaching and learning. Apart from COBUILD Project, there are many other projects aimed to apply corpus studies in foreign language teaching and learning. English Language Teaching (ELT), however, seems to be unaffected by these developments and projects in the implementation of corpus studies in language teaching because few teachers have the knowledge of corpus tools, concordance programmes and computer software programmes about corpus.

The general and specialised corpora have many direct and indirect application in ELT. Below is Figure 4, which shows the relationship between corpus and language teaching and learning.
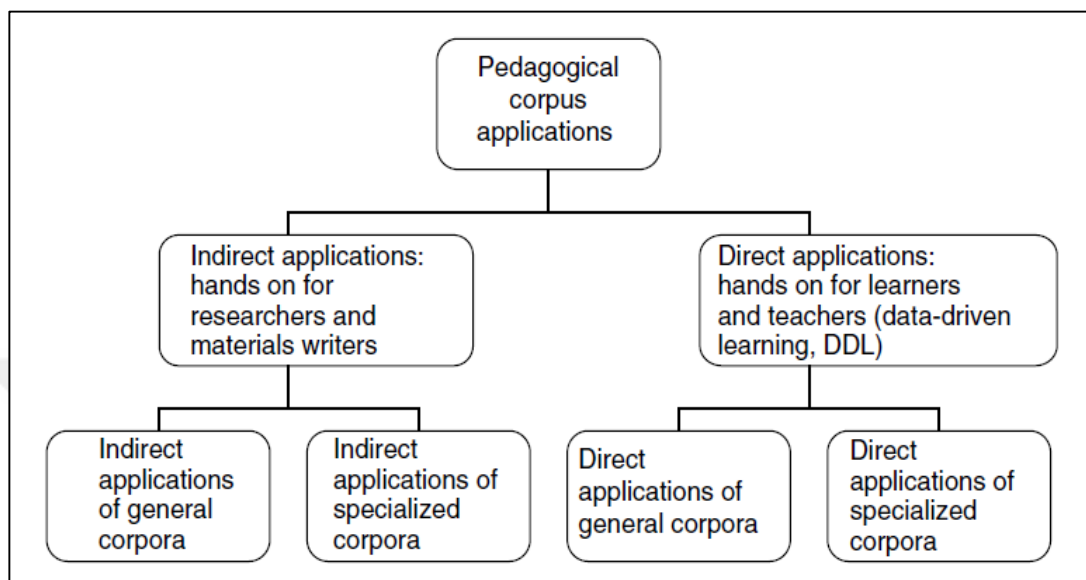
**Figure 4: The use of corpora in language learning and language teaching
(Römer, 2010:9)**

Figure 4 shows that direct applications of using corpora in language teaching and learning include the learner and the teacher who actively use corpora and concordance by means of hands on activities. The indirect applications, on the other hand, are related with researchers, material designers and syllabus makers. Römer (2010) states that direct applications of using corpora in language teaching seek the answer of *how to teach* while indirect applications seek the answer to the questions of *what* and *when to teach*.

2.4.1. Indirect Applications of Corpus in Foreign Language Teaching

The indirect applications of corpus in foreign language teaching include the preparation of coursebooks, dictionaries, grammars, material development, syllabus design and so forth. The uses of indirect corpora in foreign language teaching can be shown as in Figure 5.
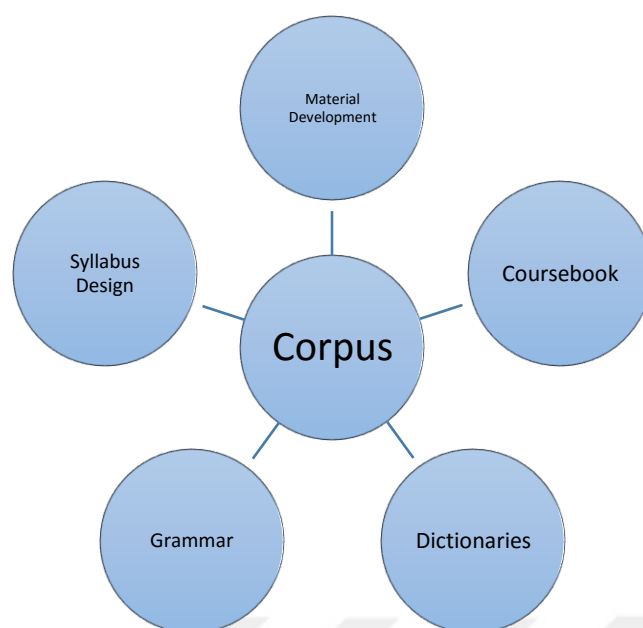
Figure 5: The indirect use of corpora in language learning and language teaching

Among the indirect uses of corpus in language teaching, the most important one is coursebook design. The coursebooks of English as a foreign language (EFL) can be examined to see what type of language data they have in presenting the language (Granger, 2002; Römer, 2004; Römer 2010). With the help of the corpus software programmes, the words or the word phrases in the coursebook can be analysed to see what the most frequent words of the coursebook are and whether the most frequent words of the coursebook correspond with the samples of the real language. The corpus software programmes also help the researchers to find which words of the coursebook collocate with each other most and whether these collocational patterns correspond with the language samples of the native speakers or not. Moreover, the semantic areas of the words in the coursebooks can be analysed by means of corpus studies. As it is seen, corpus helps the researchers to judge about the authenticity of EFL coursebooks. It is widely known that the language produced in real environments should be at the heart of the language teaching. Thus, corpus is the most useful tool in comparing the language of coursebooks with real language samples.

There are a lot of research about analysing the language of a coursebook with corpus software tools. One of the research was made by Römer (2004), who analysed EFL coursebook used in German Grammar Schools. As a result of her research, she found

that the language of the coursebooks was not genuine. Instead, it was simplified and consisted of invented samples of English. Her research also showed that the sentences in the coursebook are very artificial and may never have been used by the native speakers of English in real environments. She illustrates her research as follows with a dialogue from an EFL coursebook used in a German Grammar School (Schwarz, 1947:45) (as cited in Römer, 2004:153).

MR SNOW	:	Hello, Wendy.
MRS SNOW	:	Hello, Ron.
MR SNOW	:	Where are the girls? Are they packing?
MRS SNOW	:	Yes, they are.
MR SNOW	:	Or are they playing?
MRS SNOW	:	No, they aren't, Ron. They are packing

(Schwarz, 1947:45)

Römer (2004) sates that such a dialogue is an invented sample and does not reflect the real usage of the native speakers of English. She supports her ideas by stating that the question from the dialogue *Are they packing?* does not take place in BNC and the other question of the same dialogue *Are they playing?* takes place only once.

It is seen that encountering the invented or simplified sentences of English in the coursebooks does not help the learners to communicate in the language effectively. Thus, corpus helps researchers and teachers to judge the authenticity of the coursebook in terms of language samples, which in turn helps the language learners to gain communicative competency and be more motivated in language learning.

2.4.2. The Direct Applications of Corpus in Foreign Language Teaching

The direct applications of corpus in foreign language teaching include the pedagogical implementations of corpus in the classroom settings. Concordances, data driven learning exercises (e.g., gap filling) are examples of pedagogical implementations.

The direct applications of corpus in foreign language teaching can be carried out for several purposes. The most important purpose is that students face genuine examples of language rather than artificial ones. According to Granath (2009), the other purposes in using corpus in classroom settings are that corpus offers so many different exercises that a teacher can use in teaching, shows variation in grammar and shows the formation of syntactic arrangements of the words.

18

There are a lot of computer software programmes or internet websites that the teachers benefit from in teaching. One of them is CorpusLAB by Michael Borlow. CorpusLAB is a website and there are so many exercises about corpus that can be used by the teachers and the students. The other website is Tom Cobb's "Compleat Lexical Tutor" website. It presents vocabulary exercises, concordances and so forth. Moreover, the teachers can form texts that contain most frequent words on any subjects by using this website. Below is a screenshot of the Compleat Lexical Tutor homepage.


Figure 6: Compleat Lexical Tutor by Tom Cobb (Cobb, 2015 source: http://www.lextutor.ca/cgi-bin/range/texts/index.pl.)

Granath (2009) gives an example of an exercise that is used by the teacher in a syntax lesson in Figure 7. The purpose of the exercise is to develop the skills of the students in syntax.

1	Colin Montgomerie could find no consistency in his second **round**.
2	I turned **round** and the turret was jus a well of fire behind me. There was fire everywhere.
3	But the burglars have been **round** and stolen all the verbs and adjectives so we have to make do.
4	I had night-vision goggles on and I looked down and I could see the **round** shape of landmines.
5	And **round** every corner you could see Mount Teide itself – all 3,700m of it.
6	In the 12th **round** he was trapped on the ropes, too exhausted to escape.
7	David Starkey guides me **round** his impressive art collection in his own inimitable way.
8	I can see him clearly – slim, pale, debonair, his hair smoothed back, a college scarf **round** his neck.
9	Eventually I told him to give me back his front-door key, and phone first before coming **round**.
10	'… the infantry would come in between the tanks and the prisoners of war and **round** them up,' Brisebois said.

Figure 7: Determining the word class of round (Granath, 2009:50)

By using the exercise in Figure 7, the students may be asked to find the word classes of the word *around*, whether it is noun, adjective, verb, preposition or adverb. It is a very useful activity in that it presents real examples of language and helps the syntax development of the students.

As is seen above, if the teacher has a good knowledge of corpus, corpus tools, using corpus in language teaching and if there are computers with sufficient internet connection, corpus has an effective and unquestionable role in language teaching and learning. Bernardini (2002) agrees with that idea by stating that corpus aided language learning enhances learner motivation, fosters communicative competence and increases learner autonomy.

2.5. The Role of Vocabulary in Language Teaching and Learning

Vocabulary has an indispensable role in language learning (Evelyn and Brown, 1995; Waring and Nation; 1997; Schmitt, 2000; Thurnbury, 2002; Nation, 2006; Karatay, 2007; Nazary, 2008; Wallace, 2008; Moghadam et all, 2012; Akbari,2012; Michael et all, 2012; Toprakoğlu and Dilman, 2015). However, the importance of vocabulary has not been realized for a long time. Language teaching approaches, such as Direct Method and Audiolingualism have given more importance to acquiring grammar. The

practitioners of these approaches have stated that knowing the grammatical approaches is the core element of language teaching and learning. Many former EFL coursebooks were written to help the students to gain grammatical competence. Teachers were also in favour of teaching the grammar of the target language because they thought that grammar has a system of rules and it generated in a systematic way, which allowed the students to learn it easily. However, the practitioners of these language teaching methods, coursebook writers, teachers etc. did not take into account the idea that the aim of the language learning was not to gain grammatical rules, but to communicate efficiently in the target language. With the advent of Communicative Language Teaching (CLT), communicative competence rather than grammatical knowledge gained importance. The most important concept of CLT, that is lexical syllabus, stressed the idea that vocabulary is the core element of communication and fluency rather than accuracy should be focused on. The supporters of CLT claimed that knowing grammatical rules of the target language does not mean anything if the leaner cannot express himself/herself while, for example talking to a tourist. In this context, knowing the core vocabulary of "greeting and meeting" helps a lot and it shows that even with small number of words, the aim can be achieved.

It is understood that the role of vocabulary is seen very important in language teaching and learning and the acquisition of vocabulary should be prioritised for effective language teaching. There is a strong relationship between the acquisition of vocabulary and four skills of language (reading, listening, speaking and writing) as well as grammar.

2.5.1. Vocabulary and Reading

It is known that the unknown words that the students come across while reading a text cause problem. It shows that there is a strong connection between the vocabulary knowledge and reading comprehension. A lot of research have shown that it is necessary to know the most frequent 2000 words to understand the nearly 80% of the texts. Knowing the most frequent 2000 words helps the learners in understanding the message of the texts.

There are a lot of research that show the relationship between the knowledge of the vocabulary and reading comprehension. Coady et al. (1993) made a research to find

the effect of the knowledge of the most frequent words in English on EFL reading comprehension. The research consists of 79 students who attend university preparation programme. Two experimental group had special training on most frequent words and the control group did not have such a training. The results showed that the experimental group did better on EFL reading comprehension when compared to control group. This study shows that there is a strong correlation between the knowledge of the most frequent words and reading comprehension.

Another research was made by Zhang and Anual (2008) to find the relationship between the vocabulary and reading comprehension. Their research included 37 secondary school students who learn English in Singapore. The research showed that students who have vocabulary knowledge at the 2000-3000-word band showed better results in reading comprehension.

These studies show that there is a close relationship between vocabulary knowledge and reading comprehension. Vocabulary knowledge, especially the knowledge of the high frequent words, develops reading comprehension.

2.5.2. Vocabulary and Listening

As it is shown above, the relationship between vocabulary knowledge and reading comprehension was supported by many empirical studies. However, the relationship between vocabulary knowledge and listening comprehension has been neglected by the researchers. Nevertheless, it is clear that for effective listening comprehension, the learner needs to have enough vocabulary. Kelly (1991) explains that inadequate vocabulary prevents the EFL learners from effective listening comprehension. He further adds that the relationship between vocabulary knowledge and listening comprehension is not entirely backed by researchers. Nation (2006) states that 6000 to 7000 word families are needed to understand the 98% of the spoken discourse. He further adds that the most frequent 1000 words plus proper nouns are very important for listening comprehension in that the first 1000 words plus proper nouns make up around 85% of the spoken discourse. The relationship between listening comprehension and vocabulary knowledge was also supported by Staehr (2008) with his empirical research with 88 EFL learners from lower secondary education. He found

that the more vocabulary knowledge the learners have, the better results they have shown in listening comprehension.

Contrary to the research, Bonk (2000) found that there is a weak relationship between the vocabulary knowledge in the listening texts and listening comprehension. In the study, although some learners had less than 75% of the word families in the listening texts, they showed better results in comprehension. However, some learners could not show good results although they knew more than 90% percent of the word families in listening texts. Bonk (2000) associates these results to the capacity of the learners to use compensation strategies.

From all these studies, it can be understood that vocabulary growth helps the achievement in listening comprehension and there is a need to give emphasize to vocabulary learning for better listening comprehension in foreign language classroom.

2.5.3. Vocabulary and Writing

Much research shows that vocabulary knowledge has positive effects on receptive language skills: listening and reading. It can be said that this situation is the same with productive skills: writing and speaking. Wang (2014) states that a competent EFL student who has a rich vocabulary knowledge shows better results on any language activity and an EFL student with low vocabulary knowledge does not show the same success in language activities.

It can be said that the relationship between vocabulary knowledge and foreign language writing is complex, because writing is a process and it is composed of many variables, such as developing and formulating the ideas, planning and editing the writing, and so forth. Efficient cognitive ability and linguistic proficiency level are other elements that affect the writings of the EFL learners. As well as these variables' being important factors in effective writing, vocabulary knowledge has a huge role on effective EFL writing. A lot of research show that vocabulary richness is a determining factor on the quality of the writing. EFL students generally have difficulty in composition classes because of their limited vocabulary knowledge, and they usually use the same words in their writings which limits expressing themselves efficiently. Nation (2000) states that vocabulary knowledge is an important factor in determining

the quality of the writing. Engber (1995) states that the quality of the writing is affected by the vocabulary richness of the students. Laufer and Nation (1995) express the same ideas by stating that vocabulary size and lexical richness have a positive effect on the writings of the EFL learners.

As it is seen above, although it cannot be said that knowing too much vocabulary means being competent in writing, vocabulary knowledge is an important factor in determining the quality of the writings of the EFL learners.

2.5.4. Vocabulary and Speaking

The mechanism of human speech is very complex. It consists of many variables. Levelt (1989) describes these variables as conceptual and discursive planning and grammatical and phonological encoding. Likewise, Kormos (2006) describes three main variables in human speech. These are: conceptualization, formulation and articulation. When we speak in our mother tongue, we do not pay attention to these variables, in other words, these variables occur automatically. However, second and foreign language speakers experience greater difficulty in maintaining these processes because they do not have the same vocabulary size as the native speakers have.

It can be said that vocabulary is very important for speaking. Hilton (2008) states that very few studies have been concerned with the impact of vocabulary knowledge on real time spoken fluency because of the fact that it is not easy to analyse the speaking of an individual. Koizumi and In'nami (2013) made a research on the effect of vocabulary knowledge to the speaking of the individuals and found that speaking proficiency could be effectively predicted by vocabulary knowledge. Hilton (2008) finds the same results revealing the fundamental role of lexical competence in spoken fluency. In conclusion, these studies show that vocabulary knowledge has a central position for effective second and foreign language speaking although other variables such as syntactic, morphological, phonological knowledge and communication strategies are also important.

2.5.5. Vocabulary and Grammar

It is seen that there is a strong relationship between vocabulary and four skills. The same relationship can be seen between vocabulary knowledge and grammar. Schmitt

(2000:58) states that "lexical knowledge and grammatical knowledge are inextricably interrelated in a kind of lexico-grammar". Though it cannot be exactly said that knowing too much vocabulary helps the learners to gain grammatical competence, many studies show that knowing the most frequent words in English helps the learners' grammatical knowledge. In their study of word-class knowledge of non-native speakers of English, Odlin and Natalicio (1982) claim that the intermediate and advanced learners of English know the word classes of the most frequent words. It is seen that; the knowledge of the most frequent words helps a learner to recognize the word classes of these words.

Likewise, the grammatical knowledge may also be helpful in vocabulary learning. Schmitt (2000) states that it may be hard to acquire the less frequent words in English. Thus, presenting the word classes of the vocabulary helps the students learn this vocabulary. As well as presenting the word classes, teaching affixes to the students helps them in vocabulary retention. For example, if the student knows the meaning of the *-ful*, he/she is likely to guess the meanings of *useful* and *careful* (Schmitt, 2000).

As it is seen, not only vocabulary knowledge affects the grammatical knowledge, but also it is affected by it and vocabulary knowledge is an important factor in developing the knowledge of grammar.

2.6. Criteria for Vocabulary Selection

As it is seen, the role of vocabulary in language learning and teaching is crucial. For effective writing, speaking, listening and reading in any foreign language, the learner needs to have a good vocabulary knowledge. Because of this importance of vocabulary, the vocabulary of the coursebooks and teaching materials should be selected carefully. In this sense, vocabulary selection gains importance.

The students come across a lot of words in coursebooks, teaching materials, reading texts and so forth. However, not all of these words are equally useful. Some of the words may be out of date, old-fashioned, belong to a particular dialect, be foreign word or so forth. Thus, it is not useful to learn every word that the students encounter. It is a heavy burden for the students. These kinds of vocabulary items do not help students in an immediate environment to express themselves in the target language. Schmitt

25

(2000) states that if the student wished to survive in any short holiday in abroad, it would be enough for him to know the basic vocabulary that he/she needs like numbers from 1-10, wh questions (what, where, when, why and how) and vocabularies that include activities such as buying, asking for direction, reading signs and so forth.

As it is seen, vocabulary selection should be done in a systematic way in language teaching. Schmitt (2000) states that there are four criteria for choosing vocabulary in language teaching. These criteria are vocabulary frequency, words particularly useful in a specific topic area (technical vocabulary), vocabulary that students want to learn and vocabulary necessary for classroom management. West (1953) proposes two criteria for vocabulary selection: common words and readability. White (1989), Richards (2001) and Genç (2004) talk about six criteria while choosing vocabulary: frequency, coverage, range, availability, defining power and learnability. Ünlü (2012), likewise, proposes three criteria for vocabulary selection. These are vocabulary frequency, range and vocabulary repetition.

It is seen that, in all of these criteria, vocabulary frequency has been seen as the most important one by the researchers.

2.6.1. Vocabulary Frequency

Vocabulary frequency is the most important criterion in selecting the right vocabulary in foreign language teaching. Despite the importance of vocabulary frequency, it has been neglected by practitioners of ELT for many years and it has been gaining popularity in recent years. According to Leech (2001), there are three reasons why vocabulary frequency has been neglected in foreign language teaching. First of all, before the availability of computers, it was very hard to compile millions of words and work on them in terms of frequency. However, with the advent of computers and internet, it has become very easy for researchers to arrange and organize the language data and thus corpus studies have been applied to the field of language teaching. Thanks to the corpus software programmes, coursebooks, teaching materials and dictionaries now take into account the vocabulary frequency. For example, dictionaries, such as *Oxford Advanced Learners' Dictionary* and *Collins Cobuild Dictionary* give place to the frequencies of the words as well as their meanings. Also,

Longman Advanced American Dictionary gives frequency boxes to give information about any word's written and spoken frequency.

Figure 8: Frequency Boxes in *Longman Advanced American Dictionary* (Leech,2001)

Figure 8 shows that both *return* as a verb and *return* as a noun are very frequent in written English, however, they are not that frequent in spoken English. In Figure 8, "1" means that the word is in the most frequent first 1000 words, "2" means that the word is in the most frequent second 1000 words and "3" means that the word is in the most frequent third 1000 words. The information is very useful in deciding the usefulness of a word by the teachers and is also important for coursebook designers in selecting the right vocabulary for the coursebooks.

The second reason why vocabulary frequency has been neglected in foreign language teaching is that the practitioners of ELT gave little attention to corpus linguistics, instead, they leaned toward how people learn languages and techniques and strategies in foreign language teaching. They did not give much priority to the relationship between frequency and foreign language learning. Ellis (1994) agrees on this idea by stating that input frequency is among the factors that affects second language acquisition, however, there are very few studies to support this idea.

The third reason of the negligence of vocabulary frequency is about the difficulties of the application of it into practice. In order to decide the frequency rate of a word, many

variables arise like *range*, *coverage*, *learnability* and *communicative need* of the word, which make it a hard process (Leech,2001).

For these three reasons, vocabulary frequency has been ignored, however, in recent years, there are a lot of research about the effect of vocabulary frequency on selecting vocabulary of foreign language teaching.

Before giving information about the importance of frequency in vocabulary selection, it is beneficial to talk about how many words there are in English, how many words native speakers know and how much vocabulary a foreign language learner needs.

The first question is about the number of words in English. Waring and Nation (1997) state that it is useful to look up the largest dictionary to learn the total number of the words in English. However, this issue is a bit controversial because there are problems in determining whether the noun and the verb forms of the same word are different words. There are also problems in determining whether the compound words such as *pancakes*, *rainbow*, *jellybean* are separate words and whether the proper nouns, such as *London*, *Nutella* and *Maya the Bee* are counted as words or not. Waring and Nation (1997:6) state that;

> When compound words, archaic words, abbreviations, proper names, alternative spellings and dialect forms are excluded, and when words are classified into word families consisting of a base word, inflected forms, and transparent derivations, Webster's 3rd has a vocabulary of around 54,000 word families (Waring and Nation, 1997:6).

As it is seen, 54000 word families are beyond the dreams of the foreign language learners. The question is how many of these word families are known by the native speakers of English. The answer of this question changes according to the background, age, education level and intelligence of the native speakers. Goulden et al. (1993:341) state that although it is really hard to estimate the exact number of the words that a native speaker knows in the past, recent studies give more clear numbers on the vocabulary size of the native speakers of English. According to these studies, well-educated adult native speakers of English know approximately 17000 base words and this number shows that an individual adds two to three words to his/her memory per day. This number reaches up to 20000 for a university graduate. As it is seen, there are 54000 word families in English and only 20000 base words are known by the native

speakers of English. It is not imaginable for a foreign language learner to reach this number of words. Thus, it is very crucial for them to work with the right words instead of the useless ones. In this sense, word frequency gains importance.

Waring and Nation (1997:7) describe frequency as "how often the word occurs in normal use of the language". In this respect, some of the words in English occur very frequently and if the learner knows these words, he/she will be able to understand much of the written and spoken texts. Below is table which shows the vocabulary size and text coverage in the Brown Corpus:

Table 3: Vocabulary size and text coverage in the Brown Corpus (Francis and Kucera, 1982)

VOCABULARY SIZE	TEXT COVERAGE
1000	72.0%
2000	79.7%
3000	84.0%
4000	86.8%
5000	88.7%
6000	89.9%
15.851	97.8%

The vocabulary size in Table 3 consists of the lemmas, not word families and the table shows that by knowing 2000 lemmas in the target language, the learner reaches nearly 80% text coverage. A similar research was made by Hirsh and Nation (1992) to examine the relationship between the vocabulary size, text coverage and the density of the unknown words in English. They carried out their research on novels written for young learners.

Table 4: Vocabulary size and coverage in novels for young leaners (Hirsh and Nation, 1992)

Vocabulary Size	% Coverage	Density of Unknown Words
2000 words	90%	1 in every 10
2000 + proper nouns	93.7%	1 in every 16
2600 words	96%	1 in every 25
5000 words	98.5%	1 in every 67

Table 4 shows that the vocabulary size of the 2000 words helps the leaner to reach the 90% text coverage and it means that 9 words in 10 words are already known by the learner, which is a good level to understand the texts. It is seen that, although there are 54000 word families in English and an adult native speaker knows nearly 20000 of these words, it is enough for foreign language learners to know around 2000-3000 word families to get a text coverage of 80% in written and spoken texts. In this sense, knowing the most frequent words helps a lot to the language learner for a good comprehension.

There are a lot of research proving that high frequency words help the leaners to reach to a good percentage of text coverage. Francis and Kucera (1982) state that knowing 2000 word families helps the leaners to get nearly 80% text coverage. The study of Hirsh and Nation (2002) shows that 2000 words in novels for young leaners help the leaners to understand the 90% of the texts. Another study has been made by Schonell et al. (1956) on the oral vocabulary of adults and they found that knowing approximately 2000 word families helps the adults to understand the 99% of the spoken texts. Similar results were found by Adolphs and Schmitt (2003) who analysed modern spoken corpus (the CANCODE Corpus) and found that knowing 2000 word families make up nearly of 95% coverage. Likewise, Nation and Newton (1997) state that knowing the basic 2000 headwords makes up nearly 85% coverage in any book on any subject.

Contrary to these views, there are some researchers who state that knowing the 80% to 90% text coverage is not enough to guess the meaning of the unknown words. Research by Liu and Nation (1985) and Laufer (1989) show that at least 95% text coverage is enough to understand any spoken or written text in English. Likewise, Engels (1968) states that 95% vocabulary coverage is not enough and more vocabulary is necessary for the remaining 5% coverage to understand the texts entirely.

Despite these objections, the common thought is that knowing 2000 to 3000 most frequent words in English helps the learners to understand the majority of the texts.

It is understood from the research that frequency is an indispensable criterion to select the right vocabulary in language teaching. Limmermann and Schmitt (2005) state that if the word has a high frequency rate, then it should be taught explicitly because these

kinds of words are very valuable and develop students' ability to communicate in the target language effectively. Contrary to frequent words, low frequency words do not have a value to be taught by the teachers if they are not important for any particular context. Sinclair & Renouf (1998) and White (1988) state that frequency is the central criterion to select vocabulary for foreign language teaching coursebooks.

In this sense, vocabulary lists or word lists based on frequency gain importance for vocabulary selection for the coursebooks and teaching materials. There are many word lists which have the most frequent words in English. The most important and leading word lists are below:

2.6.1.1. General Service List (West, 1953)

The General Service List (GSL) was first developed by West in 1940s and renewed by Professor Paul Nation and his colleagues in 2010s. The New General Service List (NGSL) is based on a carefully selected 273-million-word corpus of the 1.6 billion words of Cambridge International Corpus. The 1.6 billion corpus of the Cambridge International Corpus (CIC) includes both written and spoken texts of the British and American English. Below is a table which shows the Cambridge International Corpora (CIC) used for analysis of New General Service List:

Table 5: CIC corpora used for preliminary analysis of NGSL (Browne et al., 2013)

Corpus	Running Words
Newspaper	748.391.436
Academic	260.904.352
Learner	38.219.480
Fiction	37.792.168
Journals	37.478.577
Magazines	37.329.846
Non-fiction	35.443.408
Radio	28.882.717
Spoken	27.934.806
Documents	19.017.236
TV	11.515.296
Total	**1.282.909.322**

The New General Service List presents the most important high frequency words for foreign language learners and it presents the highest coverage of written and spoken texts with fewest words (Browne et al., 2013).

2.6.1.2. The Teacher's Wordbook of 30000 words (Thorndike and Lorge, 1944)

The Teacher's Wordbook of 30000 words is based on a selected 130000 word families or 30000 lemmas of the 18-million-word written corpus. According to Thorndike and Lorge (1944), "The list tells anyone who wishes to know whether to use a word in writing, speaking, or teaching how common the word is in standard English reading matter.

2.6.1.3. American Heritage Word Frequency Book (Caroll, Davies and Richman, 1971)

American Heritage Word Frequency Book is based on a 5 million words in written texts. The words are derived from the texts used in United State Schools. The value of this wordlist comes from the fact that the texts of the wordlist come from schools. The wordlist also shows the frequency of the words according to the grade levels and subject areas.

As it is seen, wordlists are very important consultants in searching for the most frequent words in English. Waring and Nation (1997) state that wordlists play a crucial role in curriculum design of the foreign language teaching. Along with curriculum designers, coursebook writes should also refer to the wordlists while selecting the vocabulary of the language teaching coursebooks.

2.7. Common European Framework of References for Languages (CEFR)

The aim of the most teaching methods before 1970s was to focus on linguistic forms and structures. The practitioners of these teaching methods thought that if the learners gained the grammatical rules, they would acquire the linguistic competence of the target language. Nevertheless, with the late 1970s, educators started to shift from structure centred approaches to communicative ones (Widdowson, 1990; Larsen-Freeman, 2003). They thought that the students might know the rules of the grammar and linguistic structures, however, they could have difficulties in real life situations.

Communicating outside the classroom requires more than grammar rules and linguistic competence. Hymes (1972) state that knowing when to speak, how to speak and with whom to speak is the core of Communicative Language Teaching (CLT). It can be said that the advent of CLT laid the foundations of Common European Framework of Reference for Languages (CEFR) because CEFR prioritizes developing second and foreign language teaching with communicative background.

The Council of Europe announced the first version of CEFR in 1996 and after detailed evaluations and assessments, the second version of the CEFR was announced in 2001. Long before the announcement of CEFR, The Council of Europe gave importance to specifying second language learning goals and guiding the development of second and foreign language learning materials and activities by means of Communicative Approaches (Little, 2006). The Council of Europe's interest in developing second and foreign language teaching and learning started with "Threshold Level" which has a language teaching syllabus published in 1975. Threshold Level aims that the learner needs to be able to use the target language communicatively at the minimum language level. After Threshold Level, "Waystage Level" was published in 1991 and "Vantage Level" was published in 1997. These levels give great importance to learner autonomy and CLT. These three levels then were used as the description of the three stages of CEFR. After the Threshold, Waystage and Vantage levels, three more levels (Breakthrough, Effective Operational Proficiency and Mastery) were introduced by the Council of Europe, which then formed the Common European Framework of Reference for Languages.

2.7.1. What is CEFR and the need for it?

CEFR is defined by the Council of Europe (2001:1) as follows:

> The Common European Framework provides a common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, textbooks, etc. across Europe. It describes in a comprehensive way what language learners have to learn to do in order to use a language for communication and what knowledge and skills they have to develop so as to be able to act effectively. The description also covers the cultural context in which language is set. The Framework also defines levels of proficiency which allow learners' progress to be measured at each stage of learning and on a life-long basis (Council of Europe, 2001:1)

Little (2006) also defines CEFR as a composition which helps policy makers, syllabus designers, coursebook writers, researchers and practitioners of language teaching to examine foreign and second language learners' needs, define the objectives in language teaching and develop materials and assessment tools in language teaching.

CEFR is aimed to get rid of the difficulties stemming from breakdown of communication among specialists who work in the area of languages in different countries. The framework brings a common basis for the practices in the area of language teaching and presents a transparency of syllabuses, courses, objectives and assessments in modern languages (Council of Europe, 2001). Moreover, the framework fosters cooperation among European countries in the advancement of language education policy, the development of language teaching syllabus and evaluation of the language learning objectives.

The arrangement of the framework is in a taxonomic way, which starts with the simplest levels to the most complex ones and it helps us to understand the levels easily.
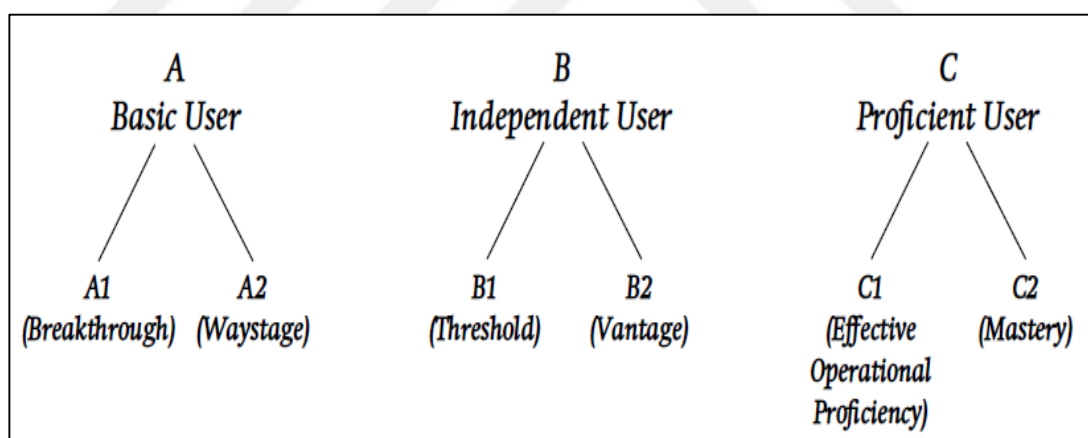
Figure 9: Common Reference Levels (Council of Europe, 2001:23)

There are three main parts in Common European Levels. These parts are A for "Basic User", B for "Independent User" and C for "Proficient User". The Basic User level has two branches: A1 (Breakthrough) and A2 (Waystage). The Independent User Level has two branches: B1 (Threshold) and B2 (Vantage). The Proficient user level has also two branches: C1 (Effective Operational Proficiency) and C2 (Mastery). The learner starts the education programme with Basic User A1 Level and he/she progresses step by step finally reaching to the proficient user level.

By means of these three main and three sub parts of the Common European Levels, the standardization in each part of education system becomes possible. For example, the curriculum design of teaching modern languages in different countries of the Europe becomes standard or the assessments of the language teaching is prepared by taking into account the Common European Levels.

Common European Levels are a general figure and the description of the six levels is given in the global scale of the CEFR. The self-assessment grid, on the other hand, presents more specific descriptors which shows the language skills of the learners at different stages.

Table 6: Common Reference Levels: Global scale (Council of Europe, 2001:24)

Proficient User	**C2**	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	**B2**	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics, which are familiar, or of personal interest. Can describe experiences and events, dreams, hopes & ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	**A2**	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Table 7: Common Reference Levels: self-assessment grid (Council of Europe, 2001:26)

		A1	A2	B1
Understanding	**Listening**	I can recognise familiar words and very basic phrases concerning myself, my family and immediate concrete surroundings when people speak slowly and clearly.	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local geography, employment). I can catch the main point in short, clear, simple messages and announcements	I can understand the main points of clear standard speech on familiar matters regularly encountered in work, school, leisure, etc. I can understand the main point of many radio or TV programmes on current affairs or topics of personal or professional interest when the delivery is relatively slow and clear.
	Reading	I can understand familiar names, words and very simple sentences, for example on notices and posters or in catalogues.	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters	I can understand texts that consist mainly of high frequency everyday or job-related language. I can understand the description of events, feelings and wishes in personal letters
Speaking	**Spoken Interaction**	I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.	I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel and current events).
	Spoken Production	I can use simple phrases and sentences to describe where I live and people I know.	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job	I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes & ambitions. I can briefly give reasons and explanations for opinions and plans. I can narrate a story or relate the plot of a book or film and describe my reactions

Writing	**Writing**	I can write a short, simple postcard, for examples sending holiday greetings. I can fill in forms with personal details, for example entering my name, nationality and address on a hotel registration form.	I can write short, simple notes and messages relating to matters in areas of immediate need. I can write a very simple personal letter, for example thanking someone for something	I can write simple connected text on topics which are familiar or of personal interest. I can write personal letters describing experiences and impressions

		B2	**C1**	**C2**
Understanding	**Listening**	I can understand extended speech and lectures and follow even complex lines of argument provided the topic is reasonably familiar. I can understand most TV news and current affairs programmes. I can understand the majority of films in standard dialect.	I can understand extended speech even when it is not clearly structured and when relationships are only implied and not signalled explicitly. I can understand television programmes and films without too much effort.	I have no difficulty in understanding any kind of spoken language, whether live or broadcast, even when delivered at fast native speed, provided I have some time to get familiar with the accent.
	Reading	I can read articles and reports concerned with contemporary problems in which the writers adopt particular stances or viewpoints. I can understand contemporary literary prose.	I can understand long and complex factual and literary texts, appreciating distinctions of style. I can understand specialised articles and longer technical instructions, even when they do not relate to my field.	I can read with ease virtually all forms of the written language, including abstract, structurally or linguistically complex texts such as manuals, specialised articles and literary works.
Speaking	**Spoken Interaction**	I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.	I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers	I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.
	Spoken Production	I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.	I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion	I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.

		I can write clear, detailed text on a wide range of subjects related to my interests. I can write an essay or report, passing on information or giving reasons in support of or against a particular point of view. I can write letters highlighting the personal significance of events and experiences.	I can express myself in clear, well-structured text, expressing points of view at some length. I can write about complex subjects in a letter, an essay or a report, underlining what I consider to be the salient issues. I can select style appropriate to the reader in mind.	I can write clear, smoothly flowing text in an appropriate style. I can write complex letters, reports or articles which present a case with an effective logical structure which helps the recipient to notice and remember significant points. I can write summaries and reviews of professional or literary works.
Writing	**Writing**			

Little (2006) states that the levels and the stages of the Common Reference Levels and self-assessment grids are descriptive. He clarifies the levels of the CEFR in four main headlines.

First, the levels of the CEFR are multidimensional. Each level reflects the principles of Communicative Approach and these principles present what the learner can do in foreign or second language. Second, the progress from one level to the other occur in the same way as in the syllabuses and the coursebooks. The levels present an acquisition process in language learning. Third, the levels of the CEFR do not present a system that grade the students. That is to say, it is possible to see, for example, B1 (Threshold), B2 (Vantage) and C1 (Effective Operational Proficiency) level learners in the same language classes. Finally, as well as cognitive ability; maturity, experience and educational advancement of the learners are necessary to reach the highest levels in CEFR (Little, 2006).

2.7.2. The A2 (Waystage) Level

The A2 Level is placed in Basic User Level of the Common Reference Levels. It has the functions, notions, grammar and vocabulary which are fundamental to achieve communicative objectives. Most of the descriptors of this level are related with social functions. Some of these functions are:

- Making and accepting offers,
- Making and responding to invitations,
- Greeting in a polite way,

- Making social exchanges,
- Discussing about doing something and going anywhere,
- Talking about work life and spare time. (Council of Europe, 2001:33)

These descriptors are the simplified versions of the descriptors of B1 Threshold Level.

Along with A2 Level, there is a Strong Waystage Level (A2+) Level which presents more active participation in communication. Some of the functions of this level are:

- Having simple face to face communication,
- Having simple exchanging of thoughts without too much effort,
- Expressing thoughts on everyday situations with the help of other speaker if necessary,
- Handling with everyday situations,
- Expressing feelings in a simple way,
- Talking about past events, daily routines, plans, likes and dislikes,
- Making comparisons about objects. (Council of Europe, 2001:34)

Below is Table 8 which shows what a language learner can do in the A2 Level of the Global Scale:

Table 8: A2 Waystage Level Global Scale (Council of Europe, 2001:24)

A2 Waystage Level	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.

As it is seen from the table, the learner at the A2 level understands the words and phrases that are frequently used. There is an attribution to the vocabulary frequency in

A2 Waystage Level in that A2 level learner needs to work with the high frequent words instead of low frequency ones. Low frequency words do not help the learners to express themselves in their immediate situations such as shopping, work place, family, personal information and the place where the learners live. Below is Table 9 which shows what a language learner can do in the A2 Level of the Self-Assessment Grid:

Table 9: A2 Waystage Level Self-Assessment Grid (Council of Europe, 2001:26)

	A2 Level
Listening	I can understand phrases and the highest frequency vocabulary related to areas of most immediate personal relevance (e.g. very basic personal and family information, shopping, local geography, employment). I can catch the main point in short, clear, simple messages and announcements
Reading	I can read very short, simple texts. I can find specific, predictable information in simple everyday material such as advertisements, prospectuses, menus and timetables and I can understand short simple personal letters
Spoken Interaction	I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.
Spoken Production	I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job
Writing	I can write short, simple notes and messages relating to matters in areas of immediate need. I can write a very simple personal letter, for example thanking someone for something

Table 9 shows that the learners develop their listening, reading, speaking (both in terms of interaction and production) and writing skills as a whole. In the listening skills, the learner is expected to catch the vocabulary of the immediate environment. In the reading skill, the learner is expected to understand the information in everyday materials, such as timetables, restaurant menus, letters, advertisements, and so forth.

In speaking skill, the learner is expected to communicate and exchange information in simple tasks and on familiar topics like family, education, job, living conditions, and so forth. In writing skill, writing about the immediate needs is prioritized. The learner, for example, can write a simple letter to thank someone.

It is clearly understood from the A2 Level Self-Assessment Grid that effective communication is the indispensable element of the language learning. All the four skills in the grid prioritize the communication of the learner. It is clear that, for effective communication, choosing the right vocabulary gains importance. The next section gives a detailed information about the vocabulary selection in A2 Waystage Level of the CEFR.

2.7.3. A2 Waystage Level and Vocabulary

Milton and Alexiou (2009) state that although the earliest developments of the CEFR gave place to the vocabulary lists (for example, the Waystage Level by Van Ek and Trim, 1990), these lists have been removed with the compilation of CEFR in 2001. Skills-based rather than knowledge-based criteria were adopted by CEFR. (Council of Europe, 2001). Nevertheless, CEFR gives information about criteria for vocabulary selection (which words are the right words in foreign and second language teaching?) and ways to develop vocabulary. Council of Europe (2001) presents a number of options in vocabulary selection as follows:

> - To select key words and phrases a) in thematic areas required for the achievement of communicative tasks relevant to learner needs, b) which embody cultural difference and/or significant values and beliefs shared by the social group(s) whose language is being learnt;
> - To follow lexico-statistical principles selecting the highest frequency words in large general word-counts or those undertaken for restricted thematic areas;
> - To select (authentic) spoken and written texts and learn/teach whatever words they contain;
> - Not to pre-plan vocabulary development, but to allow it to develop organically in response to learner demand when engaged in communicative tasks (Council of Europe, 2001:150-151).

The options of the CEFR in vocabulary selection give priority to key words and phrases, highest frequency words and the words from authentic spoken and written texts. Thus it is seen that corpus studies play a crucial role in vocabulary selection according to the CEFR. Corpus techniques present the key words and phrases and the

most frequent words with the help of the corpus software programmes. Moreover, general corpora like British National Corpus, Brown Corpus and Cambridge International Corpus have words from both written and spoken language data, which make these corpora the primary sources to apply in selection of vocabulary.

As well as wordlists, the vocabulary range in CEFR is missing in quantitative means. Milton and Alexiou (2009:196) say that "the word lists at Threshold Level (CEFR B1) contain about 2000 words and the word lists at Waystage Level (CEFR A2) contain about 1000 words". It is understood from this statement that the first versions of CEFR had the number of vocabulary size and, however, with the recent compilation of CEFR, can-do descriptors and skills are concentrated and the number of vocabulary size is excluded. Table 10 below shows the vocabulary range of the CEFR in general terms:

Table 10: Vocabulary Range of CEFR (Council of Europe, 2001:112)

	VOCABULARY RANGE
C2	Has a very good command of a very broad lexical repertoire including idiomatic expressions and colloquialisms, shows awareness of connotative levels of meaning.
C1	Has a good command of a broad lexical repertoire allowing gaps to be readily overcome with circumlocutions; little obvious searching for expressions or avoidance strategies. Good command of idiomatic expressions and colloquialisms.
B2	Has a good range of vocabulary for matters connected to his or her field and most general topics. Can vary formulation to avoid repetition, but lexical gaps can still cause hesitation and circumlocution.
B1	Has a sufficient vocabulary to express him/herself with some circumlocutions on most topics pertinent to his/her everyday life such as family, hobbies and interests, work, travel and current events.
A2	Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics.
	Has a sufficient vocabulary for the expression of basic communicative needs. Has a sufficient vocabulary for coping with simple survival needs.
A1	Has a basic vocabulary repertoire of isolated words and phrases related to particular concrete situations.

The A2 Level vocabulary range of the CEFR gives great importance to communicative needs and shows that the learner needs to have a vocabulary to deal with the simple survival needs. Although the vocabulary range of the CEFR does not show the number of the vocabulary range in the levels, because of its validity as a hierarchy, it is

understood that the learner at the A2 Waystage Level has more vocabulary than the learner at the A1 Breakthrough Level and the learner at the Proficient User Level has more vocabulary than the students at the other levels. Contrary to the general vocabulary range information in the CEFR, the researchers estimate that vocabulary range of the A2 Level is around 1500 and 2000 words (Van Ek & Trim, 1990; Meara & Milton, 2003; Milton & Alexiou, 2009). By looking at this number, it can be concluded that it is necessary to form this number with the most useful words, that is to say, the most frequent words are needed in this context.

2.8. Coursebooks as Language Teaching Sources

Ur (1996) makes the definition of the coursebook as follows, "The term coursebook is used here to mean a textbook of which the teacher and, usually, each student has a copy, and which is in principle to be followed systematically as the basis for a language course" (p. 183). Cunningsworth (1995), likewise, defines the coursebooks as sources in accomplishing the goals which have been arranged concerning the needs of the learners. It is concluded from these statements that coursebooks are among the most important sources for foreign language teaching.

Cunningsworth (1995:7) explains the roles of coursebooks in ELT as follows:

> - A resource for presentation material (spoken and written)
> - A source of activities for learner practice and communication interaction
> - A reference source for learners on grammar, vocabulary, pronunciation etc.
> - A source of stimulation and ideas for classroom language activities
> - A syllabus (where they reflect learning objectives which have already been determined)
> - A resource for self-directed learning or self-access work
> - A support for less experienced teachers who have yet to gain in confidence (Cunningsworth 1995:7).

As it is seen form the roles of the coursebooks in ELT, there needs to be an important relationship between the coursebooks and teachers in terms of achieving the goals of the foreign language teaching. According to Cunningsworth (1995:7), the aims of the coursebooks should match with the aims of the teacher, as a result of this, the needs of the language learners are fulfilled.

Kayaoğlu (2011) states that coursebook is the core of the English Language Teaching and it takes the role of the syllabus in that it regulates how the teachers teach the language and in what order the language should be taught. Thus, the analysis of the coursebook in terms of how it is designed and which kind of information it has gains importance for effective English Language Teaching.

Ur (1996) further explains the advantages of using coursebooks in foreign language teaching as follows:

- Coursebooks present a plan which help the teachers and the learners to observe the progression of the course,
- Coursebooks are the reflections of the carefully planned syllabuses, which systematize the language teaching process,
- Coursebooks have ready-made texts and learning activities which help the teachers to use their time more effectively,
- Coursebooks are not expensive teaching materials like computer software programmes and other learning tools,
- Coursebooks are easy to use, convenient to carry and it does not need electricity or computer software to be used,
- Coursebooks help the learners to gain autonomy in learning and become less dependent on teacher.

2.8.1. ELT Coursebooks in Turkish Public Schools

The practice of English Language Teaching in Turkey has gained importance along with the globalization of the world and the recent improvements in technology as well as the increasing relationship between Turkey and European Union. With this sense, Ministry of National Education in Turkey has focused on developing the practice of English Language Teaching. According to Ministry of National Education (MEB) (2013:2), "In order to continue to provide a high-quality public education for Turkey's elementary and secondary students, periodic revision of course syllabi in all subject areas is necessary to maintain an up-to date and effective curriculum that is in line with contemporary educational research and international standards for learning"

In accordance with the statements of Ministry of National Education (MEB) (2013:2), Turkish educational system was changed from 5+3 system to the 4+4+4 one. With this change, English Language Teaching started at 2nd grade rather than at the 4th grade, which, in turn, required the revision of the English Language Teaching programme.

As well as the education system, English Language Teaching programme was revised in 2013 and the syllabi of the 2nd and 3th grades became the base of the English Language Teaching (MEB, 2013).

Ministry of National Education (MEB) (2013) states that the new English Language Teaching Programme follows the principles of CEFR. Accordingly, the new programme stresses the idea of using the language in real life situations rather than emphasizing the language as a topic study. To achieve this aim, an eclectic approach (which is a mixture of different language teaching techniques and methods) has been adopted. It has been aimed that students become the active users of the language rather than solely acquiring information.

Within the new teaching programme, ELT coursebooks in Turkey have been revised by adapting the new syllabuses. The newly revised coursebooks that are used compulsorily in Turkish public primary and secondary schools are below:

Table 11: Coursebooks compulsorily used in Turkish public schools

Upturn in English 8 (for the 8th grade learners) by Mehmet Şener, Evrensel İletişim Yayınları, Ankara,2015.
Sunshine 7 (for the 7th grade learners) by Aynur Arda and Özlem Onay, Lider Yayıncılık, Ankara,2015
English Net 6 (for the 6th grade learners) by Güler Şilit, Evrensel İletişim Yayınları, Ankara,2015
English Net 5 (for the 5th grade learners) by Güler Şilit, Evrensel İletişim Yayınları, Ankara,2015
Sunshine 4 (for the 4th grade learners) by Aynur Arda and Gizem Onay, Lider Yayıncılık, Ankara, 2015

Among these coursebooks, *Upturn in English 8* has been analysed in terms of the vocabulary it has in this study since 8th grade is the step where the students have a national exam to go to the high schools and the vocabulary of the 8th grade coursebook has an important place for the national exam. Şener (2015) states that *Upturn in English 8* presents the skills in A2 Level of the of the CEFR. Şener (2015:9) further adds that:

> The course is designed for young learners and uses high-interest themes to integrate four main skills; listening, reading, speaking and writing. Vocabulary and language use are embodying in main skills. Throughout the course, students practice language in authentic situations.
> They practice pronunciation in various listening, reading and speaking activities. The course reflects the most recent approaches to language teaching and learning, and it offers up-to-date content and interesting activities in every unit. The course remains that language is best learned when it is used for meaningful communication (Şener, 2015:9).

It is claimed in this book that the four skills (reading, writing, speaking and listening) and vocabulary are integrated to the high interest themes. It is further claimed that the language is practiced in authentic situations with interesting activities and up to date content.

Nevertheless, many commercially-available coursebooks do not fit the aims and objectives of the foreign language teaching, which in turn causes failure in language teaching, especially in non-English speaking countries as Turkey where the language teaching is based mostly on coursebooks because there is not anywhere for the learners to use the target language in their daily lives. Thus, the presentation of the language in the coursebooks has an important value (Kayaoğlu, 2011).

Çakır (2010:182) criticizes the coursebooks which are taught in Turkey as follows:

> Most of the EFL coursebooks specifically designed for the students at elementary schools in Turkey are chosen by the Turkish Ministry of National Education. This ready- made materials, which unquestionably must be followed by foreign language teachers, do not always fit the needs of the learners and teachers' expectations.
> As the writers of these course materials are mostly from Turkey, they are, as expected, unable to present the authentic language in natural contexts, which unfortunately leads to teach a pure language abstaining from the original usage. This reality is not valid only for Turkey, but it is true to encounter this fact all around the world where English is taught as a foreign language (Çakır, 2010:182).

As it is seen, the presentation of the authentic language, especially the vocabulary, is very important in coursebooks for the effectiveness of the language teaching.

2.8.2. Vocabulary in EFL Coursebooks

Alcaraz (2009) states that the importance of vocabulary in foreign language teaching has been realized since 1970s with the advance of Communicative Language Teaching and is now a keystone of the language teaching. Coursebook writers are aware of this importance; thus, they give great priority to the vocabulary of the in the coursebooks.

In spite of this importance of vocabulary in foreign language teaching, the vocabulary selection in EFL coursebooks in Turkey is not based on definite basis. The English Language Teaching Programme in Turkey suggests only a few words for each unit. Below is the suggested vocabulary of the ten units of the 8th grade (A2 Level).

Table 12: Suggested Lexis for the 8th Graders in ELT Curriculum in Turkey (MEB, 2013)

Units	Suggested Vocabulary
Unit 1 (Friendship)	back up, best friend, buddy, cool, count on, get on well with somebody, go for a walk, laid-back, mate, pyjama party, secret, sharing, sleepover, slumber party, support
Unit 2 (Teen Life)	argue, casual, fashion, impressive, nerd, outfit, relationship, ridiculous, serious, snob, teenager, terrific, trendy, unbearable, types of music
Unit 3 (Cooking)	chop, dice, fry, mash, oil, pan, peel, pour, salty, slice, sour, spicy, tasty
Unit 4 (Communication)	available, connect, contact, dial, engaged, get/keep in touch, get back, hang on/up, hold line, memo, pick up, polite, put someone through
Unit 5 (The Internet)	account, attachment, browse, browser, comment, confirm, connection, delete, log on/in/off, register, reply, screen, search engine, sign in/up, social networking site, upload

Unit 6 (Adventures)	amusing, canoeing, caving, challenging, disappointing, embarrassing, entertaining, extreme, fascinating, hang-gliding, kayaking, motor-racing, rafting, skateboarding, take risks
Unit 7 (Tourism)	all-inclusive, ancient, architecture, attraction, bed and breakfast, country side, culture, destination, fascinating, historic site, incredible, resort, rural, square, urban
Unit 8 (Chores)	clean up, doing chores, iron, keep quiet, keep/break, promises, laundry, load/empty the dishwasher, make the bed, obey the rules, return books, set the table, take out the garbage/trash, tidy up, to-do list, wash/dry the dishes
Unit 9 (Science)	cell, cure, discover, explode, genius, high-tech, lab, process, result, safety, scientific, search, succeed, test tube, vaccination
Unit 10 (Natural Forces)	avalanche, disaster, drought, earthquake, flood, global warming, hurricane, land slide, melt, suffer, survivor, tornado, tsunami, volcano,

Table 12 illustrates a few words related to the units, however, it does not show the exact number of how many words each unit must have and how the vocabulary selection needs to be done. Alcaraz (2009) states that the vocabulary selection for the coursebooks should be based on two criteria: Frequency and Functionality. He further adds that frequency and functionality help the learners to communicate in English with only 4% of all the English words.

Cunningsworth (1995:41), likewise, presents criteria for vocabulary selection in coursebooks as follows;

> Is vocabulary learning material included on its own right? If so, how prominent is it? Is it central to the course or peripheral?
> • How much vocabulary is taught?
> • Is there any principled basis for the selection of vocabulary?
> • Is there any distinction between the active and passive vocabulary, or classroom vocabulary?
> • Is vocabulary presented in a structured, purposeful way? Are learners sensitized to the structures of the lexicon through vocabulary learning

exercises based on semantic relationships, formal relationships, collocations and situation-based word groups?
• Does the material enable students to expand their own vocabulary independently by helping them to develop their own learning strategies? (Cunningsworth, 1995:41)

As it is seen from the list of the Cunningsworth (1995), he prioritizes the number, usefulness, functionality, frequency and type of the words as well as appropriateness of the words to the learner and course needs in vocabulary selection.

Recent coursebooks take the criteria above into account in vocabulary selection. Thurnbury (2002) states that even though coursebooks in the past took the grammatical syllabuses into account while selecting vocabulary, nowadays, collocation and frequency are prioritized in vocabulary selections in EFL coursebooks. He gives examples of three coursebooks which claim that vocabulary selection is based on frequency, functionality and usefulness as follows;

Strong emphasis on vocabulary, with a particular focus on high frequency, useful words and phrases. (from *Cutting Edge Intermediate*)
Well-defined vocabulary syllabus plus dictionary training and pronunciation practice, including the use of phonetics (from *New Headway English Course*)
… a strongly lexical syllabus, presenting and practicing hundreds of natural expressions which students will find immediately useful. (from *Innovations*) (Thurnbury, 2002: 14)

Briefly, choosing the appropriate vocabulary for the ELT coursebooks and distributing the words in the coursebooks are very important in effective foreign language education and frequency and the role of corpora are said to have a central role in vocabulary selection for the EFL coursebooks.

CHAPTER THREE

METHODOLOGY

This study has been based on usual linguistic research methodology by applying to corpus sample. Corpus based approach has many applications on different parts of the area of linguistics, such as grammar, lexicography, language acquisition, discourse and so forth (Rayson, 2002). These applications come as a result of five main steps. These steps are;

1. Question: A research question or model is devised
2. Build: Corpus design and compilation
3. Annotate: Computational analysis of the corpus
4. Retrieve: Quantitative and qualitative analyses of the corpus
5. Interpret: Manual interpretation of the results or confirmation of the accuracy of the model (Rayson, 2002:13)

As it is seen, according to the corpus based methodology, first of all a research question or research questions are formed, then the selected corpus is designed. After that, the data in the corpus are analysed by means of software programmes. After the analysis, qualitative and quantitative results are reached and these results are interpreted. In this study, an EFL coursebook has been analysed by applying these steps.

This study has been based on the following research questions;

1. What are the most frequent words in coursebook *Upturn in English 8*?

2. How well are K1 Words (most frequent 1000 words), K2 Words (most frequent second 1000 words), academic words and Off-list words used in the coursebook *Upturn in English 8*?

3. Are there any significant differences between the corpus in the coursebook and the British National Corpus?

4. Is the corpus in the coursebook *Upturn in English 8* appropriate for A2 level according to Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR)

5. Do the most frequent words in the coursebook correspond to the interests of A2 level English Language Learners?

To answer the research questions, the coursebook *Upturn in English 8* has been analysed. Before computational analysis of the coursebook, the book has been scanned and the pages of the book have been transferred to the computer. It is important to state that the quality of the scanning should be high because the PDF file is converted to Microsoft Word file by PDF Transformer programme and this programme needs high quality scanned pages. Ünlü (2012:40) states that "the image quality of the scanning should be at least 300 dpi for obtaining more accurately read texts by the programme". With regard to this statement, the PDF file of the coursebook has been converted to a Microsoft Word file by the converter programme "Abby FineReader 12". Nevertheless, some parts of the coursebook include visuals in which the texts are not clear or hand-written, thus, these parts have not been converted to Microsoft Word file and the texts in these parts have been written to Word file manually. Below is a sample part of the coursebook which have been written to Word file manually:

Figure 10: An original extract from *Upturn in English 8*

Thus, the texts in the Unit 1 (Lesson 2, p.11), Unit 1 (Lesson 5, p.16), Unit 4 (Lesson 4, p.39), Unit 7 (Lesson 5, p.64), Unit 10 (Lesson 5, p.88) in the student's book and Unit 4 (Activity 1, p.122), Unit 4 (Activity 3, p.124) and Unit 10 (Activity 5, p.162) have been manually written by Microsoft Word programme.

After all of the texts in the coursebook have been converted to word file, the texts in the cover page, national anthem of Turkey, Atatürk's Address to the Turkish Youth, the content, plan of the student's book and references part have been excluded from the coursebook because these parts are not related with the vocabulary in the coursebook.

Rayson (2002) says that forming the frequencies of the words is a really hard process. Thus, after converting the PDF file into Word file and excluding the unrelated pages from the book, there are many steps that have to be taken. "Compleat Lexical Tutor" and "Vmatrix" are the software programmes that are used for the computational analysis of the vocabulary in the coursebook and they need TXT (rich text) format to run on the data. Thus, the texts in the Microsoft Word format have been copied and then converted to the TXT format.

Grefenstette (1999) states that the tokenisation process (setting limits to the words in the text) is the first process before running the words on the software programmes. McEnery, Piao and Xin (2000) agree with this idea by stating that demarcating the words in a text file manually or automatically decreases the error rate in computational analysis.

Bearing these thoughts in mind, first of all, the punctuation characters have been deleted from the texts (See Table 13 and Table 14).

Table 13: The original extract from the coursebook in Unit 3

Kabuli Palaw is an Afghan pilaf dish. It consists of steamed rice with lentils, carrots and lamb. It is one of the most popular dishes in Afghanistan, and it is the Afghan national dish. Afghan people make the dish with lamb, chicken, or beef. They bake it in the oven and top it with fried and sliced carrots. They also add nuts like pistachios or almonds. They put the meat in the middle of the pilaf.

Beshbarmak is the national dish of Turkmenistan. The term Beshbarmak means "five fingers", because Turkmen people eat the dish with their hands. They boil the meat with noodles and dice it with knives. They put some onion sauce on it. They usually serve it in a big round dish with lamb soup. They always say "Amin" at the end of the meal to give thanks to God.

Colcannon is a traditional Irish dish. It consists of mashed potatoes and cabbage with milk, butter, salt and pepper. It can contain other ingredients such as leeks and onions. There are many regional variations of this colcannon. Irish people often eat it with boiled meat.

Table 14: The modified text for the analysis in Unit 3

```
Kabuli Palaw is an Afghan pilaf dish It consists of steamed rice with lentils carrots and lamb It is
one of the most popular dishes in Afghanistan and it is the Afghan national dish Afghan people make
the dish with lamb chicken or beef They bake it in the oven and top it with fried and sliced carrots
They also add nuts like pistachios or almonds They put the meat in the middle of the pilaf
Beshbarmak is the national dish of Turkmenistan The term Beshbarmak means five fingers
because Turkmen people eat the dish with their hands They boil the meat with noodles and dice
it with knives They put some onion sauce on it They usually serve it in a big round dish with lamb
soup They always say Amin at the end of the meal to give thanks to God
Colcannon is a traditional Irish dish It consists of mashed potatoes and cabbage with milk butter
salt and pepper It can contain other ingredients such as leeks and onions There are many regional
variations of this colcannon Irish people often eat it with boiled meat
```

Second, the numbers have been deleted from the texts (See Table 14 and Table 15).

Table 15: The original extract from the coursebook in Unit 7

Table 16: The modified text for the analysis in unit 7

```
Listen again and answer the questions
Where is Taj Mahal What is its architectural style
How many tourists visit Niagara Falls every year What is fascinating about it
Who designed the Blue Mosque How long did its construction take
Where is Big Ben What is special about it
How tall is the Great Pyramid Why did Egyptian workers build it
```

Third, contractions, such as *he's*, *she's*, *that's* have been turned into the long forms (See Table 17 and Table 18).

Table 17: The original extract from the coursebook in Unit 9

Hello I'm Katya from Argentina and the question is "Are you interested in science?" Well, yes, I am because I'm studying environmental engineering, and right now I'm doing my thesis on glaciers, glacial melts and global warming. I'm doing that because here

Table 18: The modified text for the analysis in Unit 9

Hello I am Katya from Argentina and the question is Are you interested in science Well yes I am because I am studying environmental engineering and right now I am doing my thesis on glaciers glacial melts and global warming I am doing that because here in Argentina there are many glaciers and the global warming is accelerating glacial

Fourth, the letters which display parts, such as *A, B, C...* have been deleted because it might cause confusion whether the letter "*a*" shows the name of the part or it is an indefinite article "*a*". Furthermore, it is important to state that the full stop which indicates the end of a sentence has been deleted from the texts, however, the full stops which appear in acronyms, such as *Mr.* and *Dr.* have not been deleted and considered as one word. After all these steps have been taken, the coursebook has become ready for the computational analysis.

The computational analysis of the words in the coursebook has been performed by two different corpus software programmes. One of them is "Compleat Lexical Tutor v.8" and the other one is "Vmatrix".

3.1. Compleat Lexical Tutor v.8

Compleat Lexical Tutor is a website which has concordancers, vocabulary profiler, exercise maker about words, and so forth. It was created by Tom Cobb, who is a professor at languages at University of Quebec at Montreal. He states that Compleat Lexical Tutor has been validated empirically (T. Cobb, personal communication, November 7, 2004). The corpus programme is free and it needs Windows and Macintosh in computers to run on. Cobb (1999) states that if the learner comes across a word in different contexts, he/she can learn and use the word very easily. Thus, Compleat Lexical Tutor presents a good way of practicing vocabulary.

The programme has three parts: 1. The learners can practice vocabulary by doing online exercises. 2. Teachers can prepare exercises for their students like close tests. 3. The researchers can make studies, such as forming frequency lists and vocabulary profile of the texts, creating their own corpora, locating words and phrases in different corpora, analysing texts for repeated phrases, generating different types of statistical knowledge about vocabulary in texts and so forth. In this study, the vocabulary profile of the words in the coursebook *Upturn in English 8* has been analysed by "Compleat Lexical Tutor Version 8".

The *Vocabulary Profile* part of the "Compleat Lexical Tutor" gives information about the most frequently used 2000 words in English, Academic Word List (AWL), and Off-List Words. (See Table 19). *K1 Words* show the most frequently used 1000 words and *K2 Words* show the most frequently used second 1000 words. Within *K1 Words*, there is information about the percentage of function and content words. Moreover, the percentage of the words with Anglo-Saxon origin is displayed by means of *Vocabulary Profile* part of the programme. *AWL* shows the words that appear in academic texts. *Off-list words*, on the other hand, include proper nouns (e.g. Chicago, Henry), unusual words (e.g. Shah, zorbing), specialist vocabulary (e.g. synthetic, Richter), acronyms (e.g. Dr, Mr.), abbreviations (e.g. Lol, B4N) and misspellings.

Table 19: Vocabulary Profile of the words in the *Upturn in English 8*

	Families	Types	Tokens	Percent
K1 Words (1-1000):	709	1273	22286	75.76%
Function:	...	...	(13112)	(44.58%)
Content:	...	...	(9174)	(31.19%)
> Anglo-Sax =Not Greco-Lat/Fr Cog:	...	...	(5070)	(17.24%)
K2 Words (1001-2000):	370	530	2677	9.10%
> Anglo-Sax:	...	...	(1027)	(3.49%)
1k+2k		...	...	(84.86%)
AWL Words (academic):	141	211	895	3.04%
> Anglo-Sax:	...	...	(30)	(0.10%)
Off-List Words:	?	1032	3557	12.09%
	1220+?	3046	29415	100%

3.2. Vmatrix

Vmatrix is a software tool which performs the frequency profiling of any corpora and compares these profiles across corpora (Rayson, 2002:153). It was created by Paul Rayson, who is the director of the University Centre for Computer Corpus Research on Language (UCREL) and a reader in the School of Computing and Communication at Lancaster University in Lancaster. The software tool is not free and it is needed to make a payment for one year or two years. After the payment, a username and a password are sent to the user. Vmatrix needs an internet connection to run on.

Rayson (2002:154) states that "the Matrix method and tool assists corpus investigation by statistical comparison of frequency profiles at the lexical level and extends this to the word-class and semantic field levels". In this study, the frequency of the words has been carried out by Vmatrix (See Table 20).

Table 20: The most frequent 20 words of the coursebook *Upturn in English 8*

Word	Frequency	Relative Frequency (%)
the	1749	6.25
and	831	2.97
a	605	2.16
you	580	2.07
I	514	1.84
to	514	1.84
is	490	1.75
in	489	1.75
your	348	1.24
it	347	1.24
do	346	1.24
of	342	1.22
are	306	1.09
can	254	0.91
with	239	0.85
what	237	0.85
about	230	0.82
my	215	0.77
for	207	0.74
on	201	0.72

Table 20 shows the words in the first column, how many times they occur in the coursebook in the second column and the relative frequency values of the words within all words in the third column.

As well as the frequency of the words, word-class levels and semantic field levels of the words in the coursebook have been carried out by "Vmatrix". Moreover, the frequency of the words, word class levels and semantic field levels have been compared across British National Corpus (BNC). In order for this comparison, BNC sampler corpus has been used. The sampler of the British National Corpus has two million words chosen from the main corpus. However, the distribution of the spoken and written data has been done equally. The BNC sampler has already been analysed in terms of parts of speech tagging and semantic tagging. BNC sampler includes 2.052.440 word tokens and 53.333 word types (specific words) (Rayson, 2002:38).

In terms of comparing the frequency of the words with BNC sampler, it should be assumed that frequencies cannot be compared without normalising the values. It means that if Corpus A contains much more words than Corpus B, Corpus A naturally has more frequencies for each word. Thus, it is necessary to convert the frequency values into the percentage values as is seen in Table 21.

Table 21: Comparison of two corpora

Corpus A			Corpus B		
Word	Frequency	Relative Frequency	Word	Frequency	Relative Frequency
will	461	2.26%	will	515	1.83%

Table 21 shows the occurrences of the word "*will*" in two different corpora. The word "*will*" occurs 461 times in the Corpus A and 515 times in the Corpus B. By looking at the number of how many times the word occurs in the corpora, it would be incorrectly suggested that the word "*will*" is highly used in Corpus B. However, the observed figures should not be compared unless the values are normalised. Thus, with relative frequency of 2.26%, "*will*" is highly used in Corpus A.

While comparing two corpora, in terms of word frequency, parts of speech frequency and semantic tag frequency, the Log-likelihood value (from now on LL) is important as well as relative frequency value. Log-likelihood (LL) value represents a value for a word, parts of speech or semantic tag which has the most important relative frequency difference between two corpora (Rayson, 2002:97). Thus, the word, part of speech or semantic tag which has the most important relative frequency difference within two corpora shows the largest LL value and is placed at the top of the list. Rayson (2002) further adds that by means of LL value, the most distinctive words, parts of speech and semantic fields of one corpus can be observed as compared to the other corpus. The words, parts of speech and semantic fields which appear with similar relative frequencies within two corpora are placed at the down of the list and if they appear with different relative frequencies, they appear at the top of the list.

The words which have the most important relative frequency difference within two corpora are called key words or key items, the parts of speech are called key grammatical categories and semantic field are called key semantic fields or concepts. Granger (1993) warns that a corpus analysis is not a wholly automatic process. The investigator needs to intervene with the results and examine them qualitatively as is seen in Table 22.

Table 22: The top 10 most significant differences at word level between the coursebook *Upturn in English 8* and BNC Sampler

		Upturn in English 8		BNC Sampler			
	Word	Frequency	Rel. Freq.(%)	Frequency	Rel. Freq.(%)	O/U-use	LL
1	unit	194	0.69	38	0.00	+	1187.02
2	internet	99	0.35	0	0.00	+	710.24
3	write	136	0.49	226	0.02	+	509.15
4	sentences	63	0.23	2	0.00	+	434.22
5	your	348	1.24	3250	0.33	+	392.05
6	activity	68	0.24	29	0.00	+	371.13
7	text	63	0.23	20	0.00	+	361.43
8	the	1749	6.25	37283	3.79	+	359.71
9	photos	52	0.19	2	0.00	+	356.06
10	below	69	0.25	42	0.00	+	350.13

Table 22 shows the frequency and relative frequency for each word in the top 10. The *O/U-use* column shows overuse (+) and underuse (-) of the word in the coursebook *Upturn in English 8* with respect to the BNC Sampler. As well as most significant differences at the word level, the most significant differences at the parts of speech level and semantic field level have been formed by the Vmatrix as is seen in Table 23 and Table 24.

Table 23: The top 10 most significant differences at parts of speech level between the coursebook *Upturn in English 8* and BNC Sampler

| | Item | Upturn in English 8 | | BNC Sampler | | O/U-use | LL |
		Frequency	Rel. Freq. (%)	Frequency	Rel. Freq.(%)		
1	NN1	5188	18.55	84446	8.59	+	2325.21
2	NN2	1991	7.12	24407	2.48	+	1533.92
3	VV0	1811	6.47	26118	2.66	+	1047.66
4	APPGE	737	2.63	10384	1.06	+	445.81
5	AT	1762	6.30	38652	3.93	+	324.96
6	JJ	1554	5.55	36947	3.76	+	202.16
7	II	1570	5.61	40428	4.11	+	132.67
8	NNU	43	0.15	159	0.02	+	108.25
9	NP1	666	2.38	15229	1.55	+	103.49
10	IW	241	0.86	4230	0.43	+	89.95

Table 23 shows the frequency and relative frequency for each part of speech in the top 10. The *Item* column shows the abbreviations of the parts of speech. For example, *NN1* stands for singular common noun (e.g. book, girl), *NN2* stands for plural common noun (e.g. books, girls), *VV0* stands for root of a lexical verb (e.g. give, work), *APPGE* stands for a possessive pronoun, a pre-nominal (e.g. my, your, our), *AT* stands for article (e.g. the, no), *JJ* stands for a general adjective, *II* stands for a general preposition, *NNU* stands for a unit of measurement, neutral for number (e.g. in, cc), *NP1* stands for a singular proper noun (e.g. London, Jane, Frederick) and *IW* stands for with, without (as prepositions). Vmatrix tool presents parts of speech in 132 categories and these categories have been given in the Appendix A.

Table 24: The top 10 most significant differences at semantic field level between the coursebook *Upturn in English 8* and BNC Sampler

	Item	Upturn in English 8		BNC Sampler		O/U-use	LL
		Freq.	Rel. Freq.(%)	Freq.	Rel. Freq.(%)		
1	Q3 (Language, speech and grammar)	406	1.45	1105	0.11	+	1216.06
2	Y2 (Information technology and computing)	284	1.02	480	0.05	+	1056.13
3	Q1.2 (Paper documents and writing)	465	1.66	3398	0.35	+	686.14
4	F1 (Food)	474	1.69	3914	0.40	+	615.73
5	P1 (Education in general)	323	1.15	1928	0.20	+	574.03
6	S3.1 (Personal relationship: General)	174	0.62	752	0.08	+	395.69
7	Q2.2 (Speech acts)	478	1.71	5794	0.59	+	374.93
8	O2 (Objects generally)	392	1.40	4156	0.42	+	374.66
9	K5.1 (Sports)	169	0.60	914	0.09	+	325.74
10	Y1 (Science and technology in general)	116	0.41	415	0.04	+	298.01

Table 24 shows the frequency and relative frequency for each semantic field in the top 10. The *Item* column shows the semantic fields. Vmatrix tool presents 21 main categories and 231 sub-categories for semantic fields and these categories have been given in Appendix B.

After the computational analysis of the words in the coursebook has been performed by the software tools "Compleat Lexical Tutor v.8" and "Vmatrix", Excel software has been used to make graphics and figures. Especially the vocabulary profile of the whole words and the words in each unit have been presented as a pie chart.

The results of the computational analysis have been interpreted in terms of word, parts of speech and semantic field frequency. After this interpretation, the corpus of the coursebook *Upturn in English 8* has been examined in terms of appropriateness for A2 Level according to Common European Framework of Reference for Languages. The descriptors at A2 Level *Vocabulary Size*, *Vocabulary Range* and *Vocabulary Control* at CEFR has been taken as the basis for interpretation.

After examining appropriateness of the corpus for A2 Level of CEFR, the correspondence of the most frequent words in the coursebook to the interests of A2 Level English Language Learners has been examined. To do this, 527 English

Language Learners from different regions in Turkey have been given questionnaires to state their interest areas. The questionnaires have been sent to the different schools in seven regions in Turkey via post. The English Language Teachers have given the questionnaires to the 8th graders and then sent them back via post. After collecting 527 questionnaires, the answers of the students have been translated from Turkish to English. After the translation, the answers have been entered to the Vmatrix software tool one by one. As a result of this, semantic fields of the students' answers have been reached and these results have been compared to the semantic fields of the words in the coursebook *Upturn in English 8*.

CHAPTER FOUR

FINDINGS AND DISCUSSION

This part has been formed according to the research questions which have been stated in Chapter 1. The results of the findings have been analysed in 10 headings. These headings are: 1. Vocabulary profile of the whole words, 2. Vocabulary profile of the words in each unit, 3. The frequency of the *words* in the coursebook, 4. The frequency of the *parts of speech* in the coursebook, 5. The frequency of the *semantic fields* in the coursebook, 6. The comparison of the *frequency of the words* in the coursebook with BNC sampler, 7. The comparison of the *frequency of the parts of speech* in the coursebook with BNC sampler, 8. The comparison of the *frequency of semantic fields* in the coursebook with BNC sampler, 9. The appropriateness of the corpus in the coursebook for A2 Level of CEFR and 10. The correspondence of the corpus in the coursebook to the interests of A2 Level English Language Learners. The results have been analysed qualitatively, quantitatively and comparatively.

4.1. Vocabulary profile of the whole words in the coursebook

After the computational analysis of the words in the coursebook by "Compleat Lexical Tutor v.8", four frequency lists have been reached. These frequency lists are K1 Words which include the most frequent 1000 words in English, K2 Words which include the second most frequent 1000 words, AWL Words which include the words with lower frequency but common in academic and university texts and Off-List Words which do not belong to any of these three categories. It is necessary to state that the main aim of this study is to analyse the vocabulary in the coursebook in terms of K1+K1 Words, academic words and off-list words.
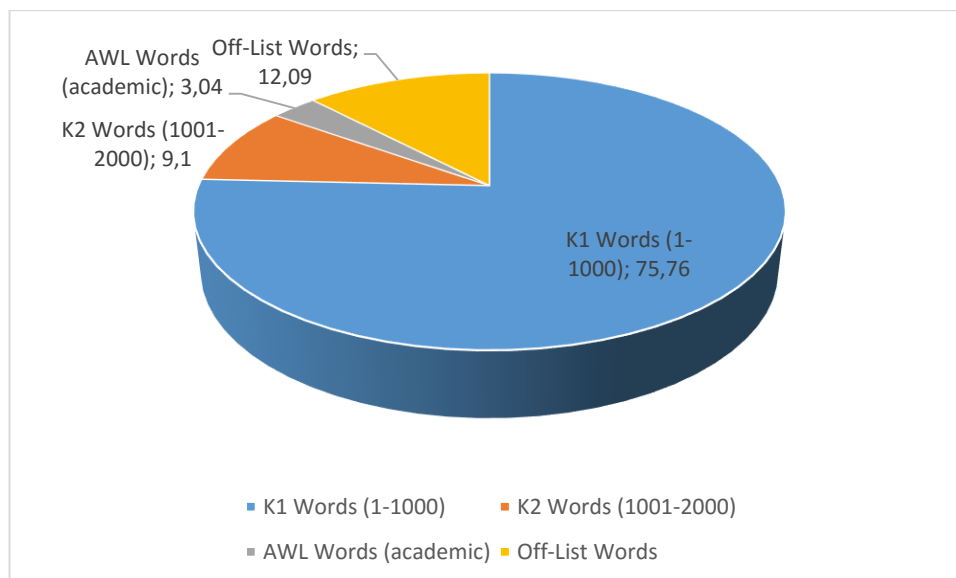
Graphic 1: Percentages of representation of frequency lists in the coursebook

Graphic 1 shows the clear domination of the K1 Words with almost 76% of the words in the coursebook belonging to this class. This is followed by Off-List Words with 12.09% and K2 Words with 9,1%. AWL Words List has the least percentage with 3,04%. It is clearly seen from the graphic that K1 Words and K2 Words form almost 85% of the words in the coursebook. However, it should be kept in mind that K1 Words is generally made up of function words as is seen in Table 25.

Table 25: Vocabulary Profile of the words in *Upturn in English 8*

	Families	**Types**	**Tokens**	**Percent**
K1 Words (1-1000):	709	1273	22286	75.76%
Function:	...	...	(13112)	(44.58%)
Content:	...	...	(9174)	(31.19%)
> Anglo-Sax =Not Greco-Lat/Fr Cog:	...	...	(5070)	(17.24%)
K2 Words (1001-2000):	370	530	2677	9.10%
> Anglo-Sax:	...	...	(1027)	(3.49%)
1k+2k		...	...	(84.86%)
AWL Words (academic):	141	211	895	3.04%
> Anglo-Sax:	...	...	(30)	(0.10%)
Off-List Words:	?	1032	3557	12.09%
	1220+?	3046	29415	100%

It is seen from Table 25 that 44.58% of the K1 words are made up of function words: prepositions (of, at, in…), pronouns (he, they, anybody…), determiners (the, a, my…), conjunctions (and, that, when…), auxiliary verbs (am, is are, have…) and particles (no, not, nor…). Only 31.19% of the K1 Words are made up of content words like nouns, verbs, adjectives and adverbs. Thus, it is quite natural that the percentage of the K1 Words are high in the coursebook because of the high percentage of the content words. In this context, the words in the coursebook correspond with the 1000 most frequent words in General English.

The same conclusion cannot be made with the K2 Words. K2 Words form only 9.10% of the whole words in the coursebook. K2 Words are made up of content words and are more essential for the meaning than K1 Words. Thus it is clear that there should be an increase in the percentage of second 1000 most frequent words. The 8th grade English Language Learners in Turkey have been learning English for six years, therefore, it is assumed that most of the K1 Words (1-1000) have been learned or handled to a certain extent. In this regard, more prominence needs to be given to increasing the number of the K2 Words (1001-2000) within the coursebook.

Another surprising result is that the percentage of the Off-List words is unusually high, contributing 12.09% to the total. Off-list words may be proper nouns, uncommon words, abbreviations, acronyms, words belonging to different dialects, special vocabulary, misspellings and so forth. However, some of the words such as *amazing, bye, chat, dentist, e-mail, festival, guitar, jogging, mall, notebook, ok, photos, honey, hiking, tablet, teenager, theatre* and *upload* are classified as Off-List Words despite their being very functional for communicative purposes. (See Appendix C).

The least percentage of words belongs to AWL Words with 3.04%. It can be stated that the vocabulary of the coursebook includes very few academic words, which do not cause much difficulty for the students.

In conclusion, it can be said that the percentage of the K1 Words and Academic Words are accepted in number, both taking the students' needs into account. However, the percentage of K2 Words is very low and there should be an enormous increase in the number of these words because this list includes content words and they are highly functional in the communicative context of the students. In contrast, the percentage of

Off-List Words is very high, which causes problems for vocabulary retention of the learners because dealing with Off-List Words is time consuming and not meaningful. Thus, there needs to be a decrease in the percentage of the OFF-List Words.

4.2. Vocabulary profile of the words in each unit

The coursebook *Upturn in English 8* has ten units. The vocabulary of each unit has been analysed and compared with one another in terms of four criteria: 1. K1 Words (1-1000), K2 Words (1001-2000), AWL Words and Off-List Words.
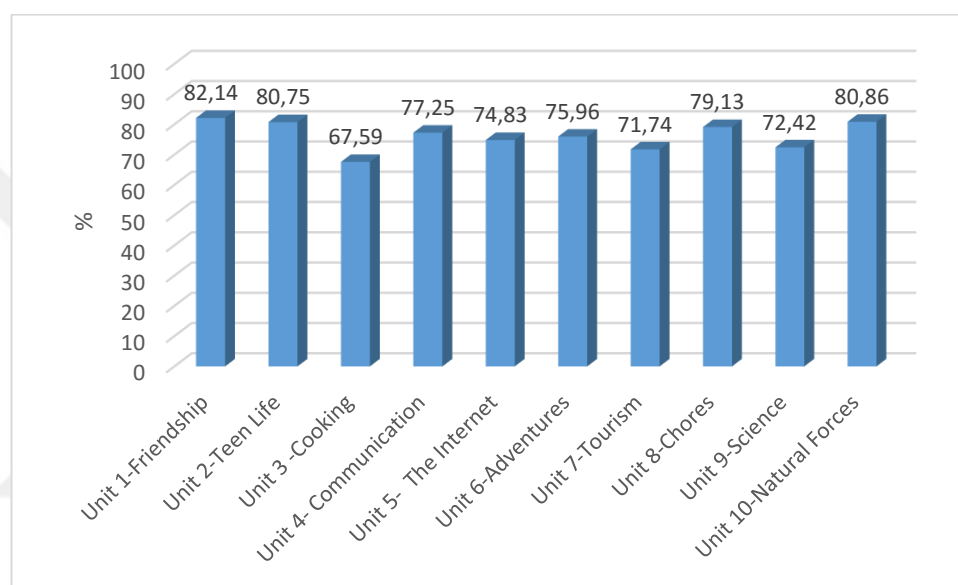
Graphic 2: Analysis of the vocabulary items of the ten units in terms of K1 Words (1-1000)

It is seen from Graphic 2 that Unit "Friendship" has the most K1 words with the percentage of 82.14%. It is followed by Unit "Natural Forces" with 80.86% and then comes Unit "Teen Life" with 80.75%. The last three units which do not have enough 1000 most frequent words in General English are Unit "Science" with 72.42%, Unit "Tourism" with 71.71% and Unit "Cooking" with 67.59%. Unit "Cooking" has the lowest number of K1 Words, which can be regarded low. It may because of the fact that there are many terms about cooking and measurement. An increase in the number of K1 words is necessary for this unit.
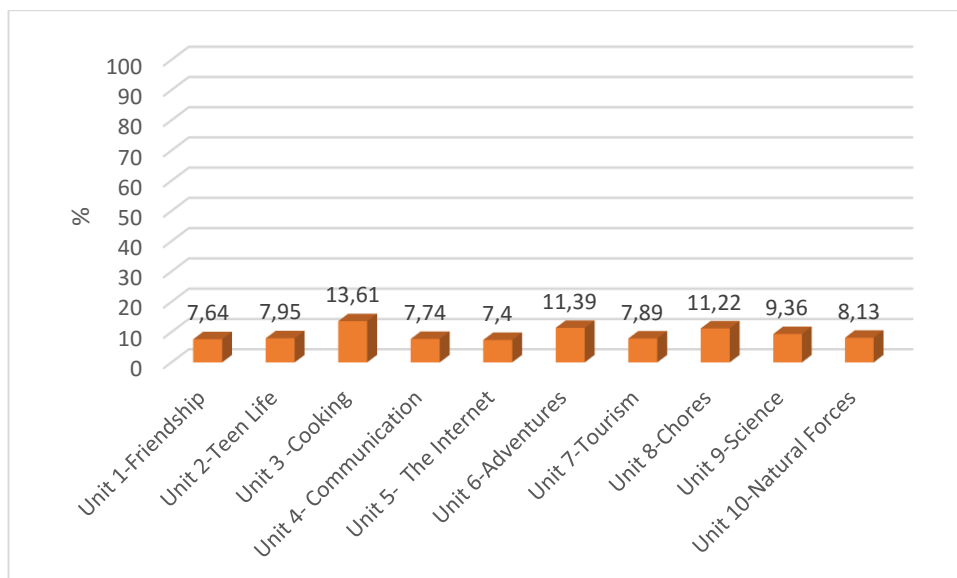
Graphic 3: Analysis of the vocabulary of the ten units in terms of K2 Words (1001-2000)

K2 Words consist of content words, which is very vital for the communicative competence of the students, thus highest number of second 1000 most frequent words in the units is useful. Graphic 3 shows that Unit "Cooking" has the most K2 Words with the percentage of 13,61%. It is followed by the Unit "Adventures" with 11.39% and then Unit "Chores" with 11.22%. The units which have the low percentage of K2 Words are "Friendship", "Teen Life", "Communication", "Internet" and "Tourism". It is remarkable to state that although Unit "Cooking" has the lowest number of K1 Words, the same unit has the highest number of K2 Words.
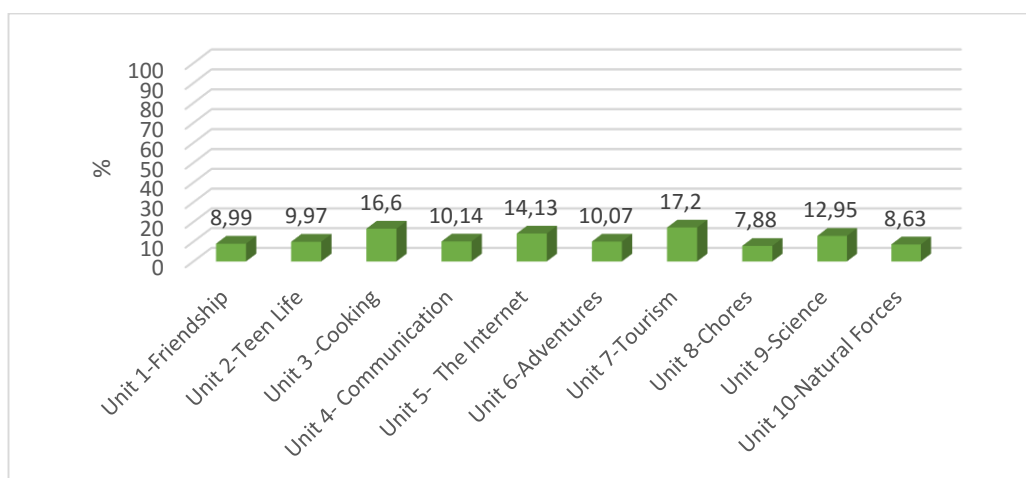
Graphic 4: Analysis of the vocabulary of ten units in terms of Off-List Words

It is clearly seen from Graphic 4 that the percentages of the Off-List Words in the coursebook is extremely high. Especially Unit "Tourism" with 17,2%, Unit "Cooking" with 16,6%, Unit "Internet" with 14,13% and Unit "Science" with 12,95% have too many Off-List Words and these unusually high occurrences should be taken into consideration because it effects the vocabulary learning and retention of the students in a negative way.

Graphic 3 and Graphic 4 also show that except form units "Adventure" and "Chores", the number of the Off-List words is higher than the number of the K2 Words in the other units, which is unacceptable because working with Off-List Words in the classroom is time consuming. When the Off-List Words in the coursebook are observed, it is clearly seen that a number of words are proper nouns: *Ally, Amanda, Ashley, Birmingham, Darren, Dave, Ellen, Frank, George, Glen, Helen, Henry, Hillary, Jack, Jane, Kate, Kevin, Kahn, Liz, Mary, Mike, Nancy, Nicole, Pete, Sally, Sam, Samuel, Sonya* and Suzy are some of the proper nouns in the first unit. *Canada, Clare, Daniela, Elmo, Italy, Kevin, Megan* and *Taylor* are some of the proper nouns in the second unit. *Afghan, Beshparmak, Colcannon, Irish, Kabuli, Lamington, Pavlova, Turkmen* and *Zealand* are some of the proper nouns in the third unit. *Albert, Alice, Amanda, Betty, Chinese, Essex, Eric, French, Harfield, Jill, John, Linda, Parker, Martin, Peterson, Racheal, Rick, Rosa, Sanders, Scott, Steve, Victor* and *Yukon* are the proper nouns in the fourth unit. *Africa, America, Asia, Bob, Europe, French, Helen, Jason* and *Ken* are the proper nouns in the fifth unit. *Italian, Joe, Turkey, USA* are some of the proper nouns in the sixth unit. *American, Ana, Anatolia, Antony, Arab, Byzantines, Canada, Caribbean, Cilicia, Cleopatra, Colorado, Dubai, Egyptian, Ikbebet, Harire, Himeji, Hittites, Jamaica, Kingston, Kremlin, Madrassas, Mesopotamia, Moscow, Iskender, Romans Russia, Sumela, Tarsus, Unesco* and *Zigana* are some of the proper nouns that belong to the seventh unit. *Belinda, Betsy, Curtis, Diana* and *Jeremy* are some of the proper nouns in the eighth unit. *Alexander, Archimedes, Babbage, Benjamin, Charles, Edison, Franklin, Galileo, Hadzicki, Harvard, Maya, McClintock, Soviet, Tereshkova* and *Watson* are some of the proper nouns that belong to the ninth unit. And finally, *Chili, Oklahoma, Taiwan, Xinhua* and *Mexico* are some of the proper nouns in the tenth unit.

Apart from proper nouns, there are technical terms in the units such as *brake, gear, lever, mudguard, delete, scan, processor, reset, antenna, heritage, astronomy, centimetres, cells, diagnose, micro, polonium, vaccine, ultra, cosmos, drug, algorithm, onda, organisms, drought, flash flood, hurricane, radar, Richter, tornado* and so forth.

Moreover, the words that belong to different dialects and languages, such as *Bonjour, Guten, Merhab*a, abbreviations such as *HAND (have a nice day), B4N (bye for now), CUL8R (see you later), 2NITE (tonight), LOL (lots of love) and BBL (be back later)*, acronyms such as *sms, vga, dvd, dna, lab, etc*, are the words in the units that belong to Off-List Words category.

Off-List Words' being higher in the number in the coursebook effects the vocabulary learning in a negative way. Thus, the number of the Off-List Words should be restricted and more prominence should be given to K2 Words especially for the units "Tourism", "Cooking", "Internet" and "Science".
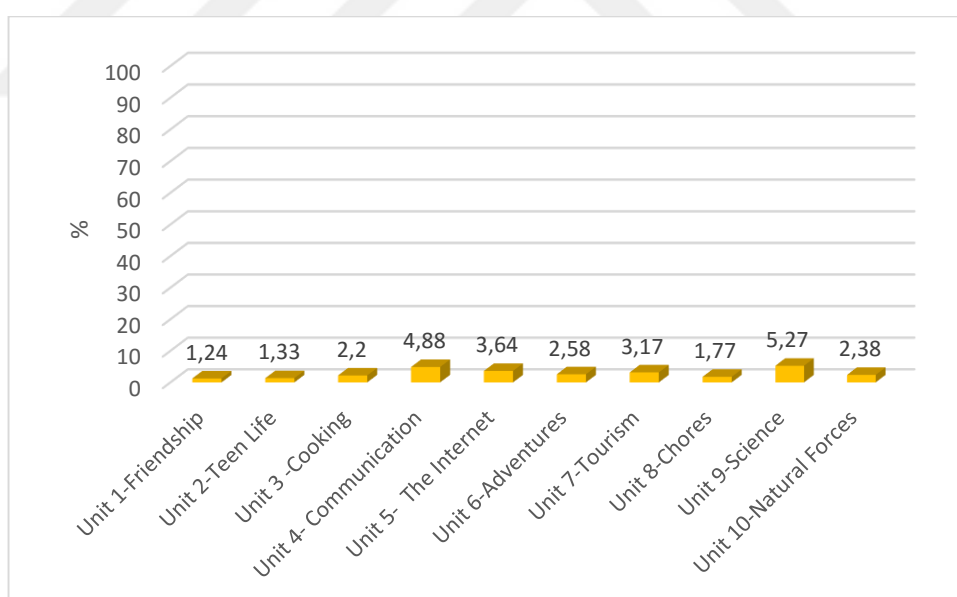
Graphic 5: Analysis of the vocabulary of ten units in terms of AWL Words

It is seen from Graphic 5 that Unit "Science" has the most academic words with the percentage of 5.27%. It is followed by Unit "Communication" with 4.88% and Unit "Tourism" with 3.17%. The units "Friendship", "Teen Life", and "Chores" have the least number of academic words, which have a positive effect on vocabulary learning

of the students. It is clearly seen that the number of the academic words in the fourth, fifth and ninth unit should be decreased because these words do not benefit to the immediate communicative environment of the students.

In conclusion, while vocabulary in some of the units in the coursebook corresponds with the most frequent 2000 words in General English, the vocabulary in some of the units does not correspond as is seen in Graphic 6.
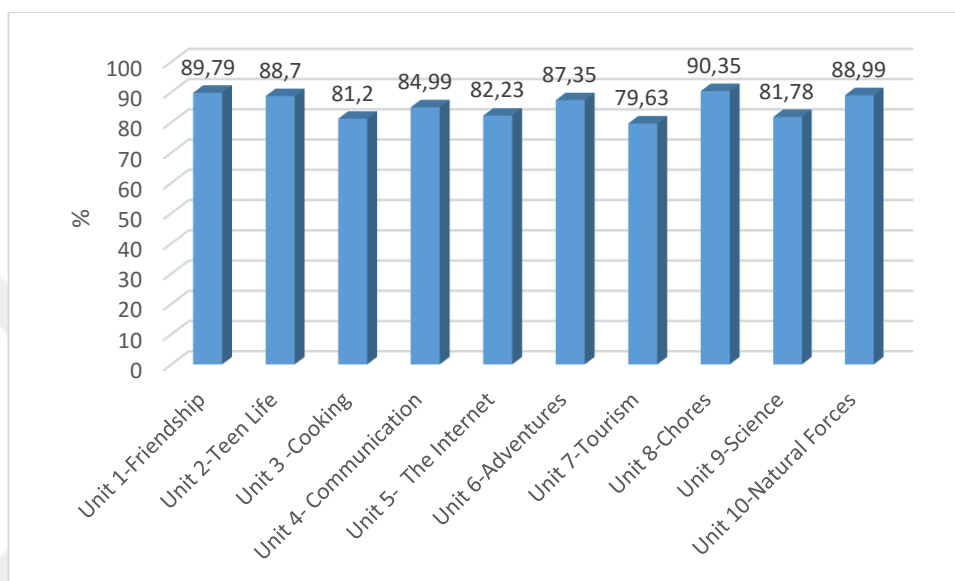
Graphic 6: Analysis of the vocabulary of the ten units in terms of K1+K2 Words (1-2000)

Graphic 6 shows that Unit "Chores" represents the most frequent 2000 words best with the percentage of 90,35% and takes the students' needs into account in terms of the vocabulary profile. However, other units, especially units "Cooking", "The Internet", "Tourism" and "Science", should be taken into consideration because of the low percentage of K1+K2 Words and a revision is needed in terms of vocabulary choice for these units.

4.3. The frequency of the words in the coursebook

After the analysis of the vocabulary in the coursebook in terms of vocabulary profile, that is to say, whether the most frequent 2000 words are dominant in the coursebook or academic words and off-list words dominate the vocabulary of the coursebook, the corpus of the book has been analysed in terms of the frequency of the words.

70

The frequency list tells us information about the core vocabulary of the corpus in the coursebook. Below is Table 26 which shows the most frequently used 100 words in the coursebook. The frequency of the whole words in the coursebook are given in Appendix D.

Table 26: Most Frequent 100 words in *Upturn in English 8*

Word	Frequency	Relative Frequency (%)
the	1749	6.25
and	831	2.97
a	605	2.16
you	580	2.07
I	514	1.84
to	514	1.84
is	490	1.75
in	489	1.75
your	348	1.24
it	347	1.24
do	346	1.24
of	342	1.22
are	306	1.09
can	254	0.91
with	239	0.85
what	237	0.85
about	230	0.82
my	215	0.77
for	207	0.74
on	201	0.72
unit	194	0.69
like	151	0.54
at	143	0.51
write	136	0.49
not	128	0.46
have	123	0.44
or	116	0.41
how	116	0.41
they	108	0.39
use	107	0.38
internet	99	0.35
we	96	0.34

Word	Frequency	Relative Frequency (%)
am	94	0.34
read	94	0.34
an	94	0.34
he	87	0.31
would	83	0.30
make	83	0.30
going to	80	0.29
that	80	0.29
people	79	0.28
questions	76	0.27
there	71	0.25
when	70	0.25
below	69	0.25
but	68	0.24
activity	68	0.24
she	68	0.24
information	67	0.24
check	66	0.24
go	65	0.23
usually	65	0.23
school	64	0.23
sentences	63	0.23
text	63	0.23
lesson	60	0.21
this	59	0.21
s	59	0.21
be	58	0.21
answer	58	0.21
share	57	0.20
ask	55	0.20
simple	55	0.20
friends	54	0.19
does	53	0.19
help	53	0.19
example	52	0.19
very	52	0.19
photos	52	0.19
science	52	0.19
then	50	0.18
her	50	0.18

Word	Frequency	Relative Frequency (%)
if	50	0.18
because	49	0.18
more	49	0.18
water	49	0.18
well	48	0.17
some	48	0.17
take	47	0.17
many	47	0.17
doing	46	0.16
them	46	0.16
why	46	0.16
understand	45	0.16
listen	44	0.16
put	44	0.16
please	44	0.16
one	43	0.15
short	43	0.15
first	43	0.15
so	42	0.15
world	42	0.15
his	42	0.15
will	42	0.15
work	41	0.15
sports	41	0.15
hello	41	0.15
who	41	0.15
prefer	41	0.15
than	41	0.15

As it is seen from Table 26, generally the prepositions, pronouns, determiners, conjunctions, auxiliary verbs and particles are on the top of the list. It can be said that their being on the top of the list results from their functions; in other words, these words are not repeated intentionally.

It is seen that the definite article "*the*" is the most frequent word in the corpus with relative frequency of 6.25%. It is followed by conjunction "*and*" with relative frequency of 2.97%, indefinite article "*a*" with relative frequency of 2.16 and personal pronoun "you" with relative frequency of 2.07%.

O'Keeffe, McCarthy and Carter (2007) state that if the personal pronouns "*I*" and "*you*" and the discourse markers such as *yeah*, *ok*, *hello*, *er* are reflected in the high rank of the frequency list of any corpus, it can be said that the corpus presents a speaker-listener environment, which is beneficial for the communicative environment of the students. O'Keeffe, McCarthy and Carter (2007) further add that if there is a predominance of the third person references with prepositions and conjunctions in any corpus, the corpus represents the "*the world out there*". These kind of words are functional rather than lexical, they have little help to the content of the corpus. With this respect, it is seen from the most frequent 100 words in the coursebook that although the personal pronouns "*I*" and "*you*" are on the top of the list, the discourse-marking items are not among the most frequent words in the list. Only the word "*Hello*" is in the list of most frequent 100 words. This shows that the vocabulary of the coursebook does not help the communicative environment of the students.

In section 4.6, the frequencies of the words have been compared with the BNC and the words which are unusually highly repeated or underused with respect to the BNC have been analysed.

4.4. The frequency of the parts of speech in the coursebook

As it is stated before, Vmatrix software tool analyses the corpus of any text in terms of 132 parts of speech categories. The analysis gives both the frequencies and the relative frequencies of the parts of speech of any corpus. Below is Table 27 which shows the frequency of parts of speech of the words in the coursebook.

Table 27: The frequency of *parts of speech* of the words in *Upturn in English*

POS	Freq.	Rel. Freq. (%)	POS	Freq.	Rel. Freq. (%)
NN1	5188	18.55	MC	108	0.39
NN2	1991	7.12	NN	108	0.39
VV0	1811	6.47	PPIS2	96	0.34
AT	1762	6.30	VDI	95	0.34
II	1570	5.61	VBM	94	0.34
JJ	1554	5.55	VH0	90	0.32
CC	949	3.39	RG	85	0.30
VVI	946	3.38	NNT2	78	0.28
APPGE	737	2.63	DD	75	0.27
AT1	703	2.51	PN1	74	0.26
NP1	666	2.38	MD	69	0.25
RR	598	2.14	CCB	68	0.24
PPY	590	2.11	VVGK	66	0.24
PPIS1	529	1.89	VBI	60	0.21
VBZ	490	1.75	DA2	59	0.21
VVG	445	1.59	VDZ	53	0.19
VM	433	1.55	CST	51	0.18
IO	344	1.23	NPD1	50	0.18
PPH1	342	1.22	VDG	48	0.17
TO	309	1.10	PPHO2	46	0.16
VBR	306	1.09	DB	43	0.15
VD0	264	0.94	NNU	43	0.15
DDQ	256	0.92	NNO	43	0.15
IW	241	0.86	CSN	41	0.15
VVD	211	0.75	VHI	40	0.14
IF	209	0.75	II21	40	0.14
CS	196	0.70	PPIO1	39	0.14
RRQ	192	0.69	RR21	39	0.14
RT	182	0.65	RGQ	39	0.14
VVZ	173	0.62	PNQS	38	0.14
VVN	169	0.60	VBDZ	38	0.14
PPHS1	155	0.55	MC1	35	0.13
NNT1	153	0.55	DD2	32	0.11
RL	152	0.54	EX	32	0.11
UH	140	0.50	PPHO1	31	0.11
XX	135	0.48	VHZ	31	0.11
ZZ1	113	0.40	JJT	30	0.11
PPHS2	109	0.39	RGR	28	0.10
DD1	109	0.39	JJR	27	0.10
			RP	26	0.09
			CSA	26	0.09

POS	Freq.	Rel. Freq. (%)
PPIO2	24	0.09
RGT	24	0.09
RRR	23	0.08
VDD	22	0.08
DAR	17	0.06
NNB	15	0.05
FO	15	0.05
RRT	14	0.05
PPX1	13	0.05
VHG	12	0.04
ND1	12	0.04
RR22	11	0.04
VBDR	11	0.04
NNU2	11	0.04
NPM1	8	0.03
DA	8	0.03
II31	8	0.03
PN	8	0.03
DA1	7	0.03
RA	6	0.02
VHD	6	0.02
DAT	5	0.02
NPD2	5	0.02
NNL1	4	0.01
VBG	4	0.01
NNU1	4	0.01
CS31	4	0.01
PPX2	3	0.01
REX21	3	0.01
NN121	3	0.01
CSW	3	0.01
DB2	3	0.01
GE	3	0.01
PPX221	2	0.01
NP2	2	0.01
VDN	2	0.01
NNO2	2	0.01
RRQV	2	0.01
PPGE	1	0.00

JJ21	1	0.00
JJ22	1	0.00
JK	1	0.00
RG21	1	0.00
RG22	1	0.00
FW	1	0.00
DDQV	1	0.00
VHN	1	0.00
NNL2	1	0.00

The POS column in Table 27 shows the abbreviations of the parts of speech. The explanations of these abbreviations have been given in Appendix A. Below is Graphic 7, which shows the most frequently used 21 parts of speech of the words in the coursebook and which have the relative frequencies over 1.00
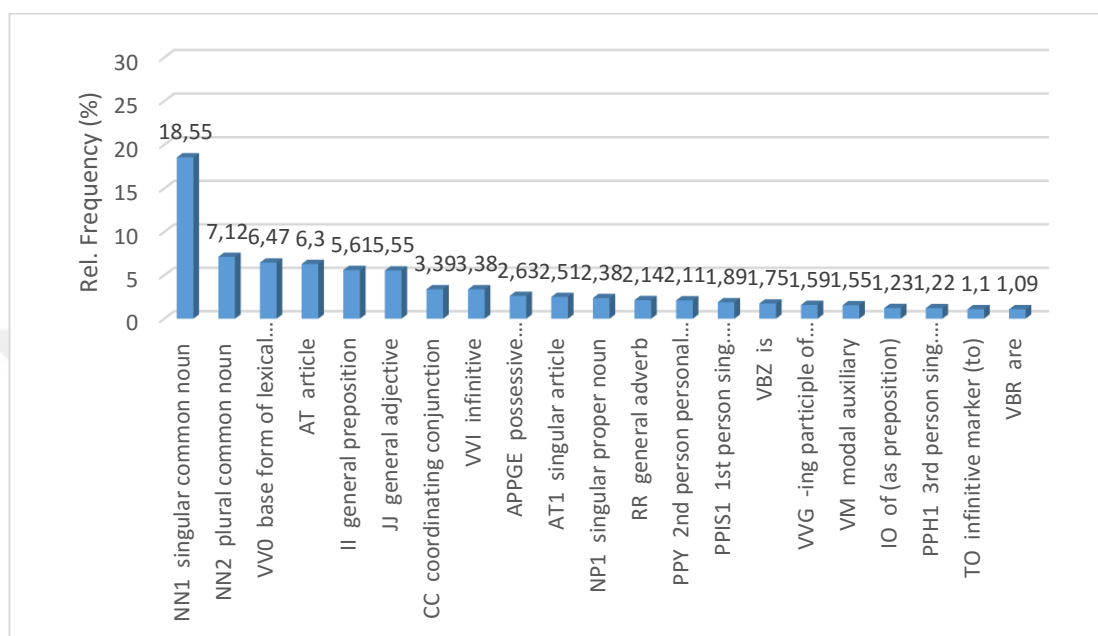
Graphic 7: The most frequently used 21 parts of speech of the words *in Upturn in English 8*

The most frequently used part of speech of the words in the book is "*Singular Common Noun*" (e.g. book, girl) with the relative frequency of 18.55%. It is followed by "*Plural Common Noun*" (e.g. books, girls) with the relative frequency 7.12% and then comes "*Base Form of Lexical Verb*" (e.g. give, work) with relative frequency of 6.47%. Graphic 7 shows the clear dominance of singular common nouns throughout the words of the coursebook.

Another interesting result from the frequency of the parts of speech of the words is personal pronouns. Graphic 9 shows the frequencies of the personal pronouns in the coursebook.
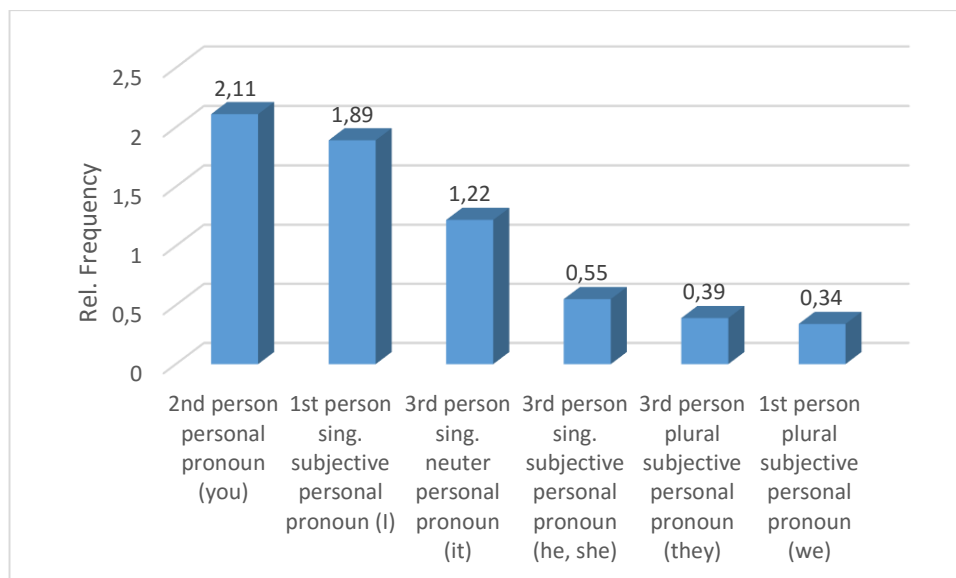
Graphic 8: The frequencies of personal pronouns in *Upturn in English 8*

Graphic 8 shows the dominance of the 2nd personal pronoun "*you*" and 1st personal singular subjective personal pronoun "*I*" over the other personal pronouns. However, these personal pronouns have not been supported by discourse markers, such as *well*, *right*, *yeah* on the top of the frequency list, therefore, they do not present important communicative functions.

The 3rd person singular subjective personal pronouns "*he*" and "*she*" have the relative frequency of 0.55%. Between these pronouns, the pronoun "*he*" has been used 87 times in the coursebook with relative frequency of 0.32%. Contrary, the other pronoun "*she*" has been used 68 times with the relative frequency of 0.24%. This shows that there is a masculine dominance on the usage of the 3rd person singular subjective personal pronouns.

When the most frequently used parts of speech of the words are analysed in a more detailed way, it is seen that singular nouns outnumber the plural nouns in the corpus of the coursebook. For example, while singular common nouns (e.g. book, girl) have relative frequency of 18.55%, plural common nouns (e.g. books, girls) have the relative frequency of 7.12%. Likewise, singular proper nouns (e.g. London, Jane, Frederick) have the relative frequency of 2.38% and plural proper nouns (e.g. Browns, Reagans, Koreas) have the relative frequency of 0.01%. Singular determiners (e.g. this, that, another) have the relative frequency of 0,39%, while plural determiners (these,

78

those) have the relative frequency of 0.11%. Singular weekday nouns (e.g. Sunday) have the relative frequency of 0.18%, while plural weekday noun (e.g. Sundays) have the relative frequency of 0.02%. Temporal singular nouns (e.g. day, week, year) have the relative frequency of 0.55% and temporal plural nouns (e.g. days, weeks, years) have the relative frequency of 0.28%. Interestingly, while the singular month nouns (e.g. October) are used 8 times in the coursebook with the relative frequency of 0.03%, plural month nouns (e.g. Octobers) have not been used in the coursebook. All these examples show the dominance of singular nouns over plural ones.

Another analysis of Table 27 shows that the definite article "*the*" has been used more than the indefinite articles "*a*" and "*an*". While the definite article "*the*" has been used 1749 times in the coursebook with the relative frequency of 6.25%, the indefinite article "*a*" has been used 605 times with the relative frequency of 2.16% and the other definite article "*an*" has been used 94 times with the relative frequency of 0,34%. The definite article is used two times higher than the indefinite articles. This result shows that definiteness plays an important role on the vocabulary in the coursebook.

Apart from the definiteness, it is seen from Table 27 that the usage of possessive pronoun, pre-nominal (e.g. my, your,) is very high in the coursebook. The possessive pronouns, prenominal have been used 737 times with the relative frequency of 2.63%. This high use of possessiveness is also observed in classroom. For example, students generally say "my teacher" instead of "teacher" in the classes. In part 4.7., the use of possessiveness has been compared to BNC to see whether it has been over or underused with respect to British National Corpus.

4.5. The frequency of the semantic fields in the coursebook

It has been stated in the Methodology part that Vmatrix software tool analyses the corpus of any text in terms of 21 main and 231sub-categories in terms of semantic fields. The analysis gives both the frequencies and the relative frequencies of the parts of speech of any corpus. Table 28 below shows the frequency of semantic fields of the words in the coursebook.

Table 28: The frequency of *semantic fields* of the words in *Upturn in English 8*

Semtag	Freq.	Relative Frequency (%)
Z5	7321	26.17
Z8	3109	11.11
A3+	788	2.82
A1.1.1	645	2.31
Z1	499	1.78
Q2.2	478	1.71
F1	474	1.69
Q1.2	465	1.66
M1	420	1.50
Q3	406	1.45
A7+	398	1.42
O2	392	1.40
P1	323	1.15
Z4	294	1.05
T1.3	288	1.03
Q2.1	285	1.02
Y2	284	1.02
M6	278	0.99
A9+	275	0.98
S5+	250	0.89
Z2	240	0.86
Z99	204	0.73
E2+	200	0.71
A2.2	200	0.71
A5.1+	191	0.68
S4	191	0.68
M2	188	0.67
S3.1	174	0.62
K5.1	169	0.60
T1.1.3	163	0.58
M7	161	0.58
A1.5.1	153	0.55
Z6	148	0.53
S2	144	0.51
W3	142	0.51
N5	141	0.50
N4	137	0.49
A6.2+	136	0.49

Semtag	Freq.	Relative Frequency (%)
X2.4	135	0.48
N1	129	0.46
X2.2+	128	0.46
M3	122	0.44
A13.3	121	0.43
A4.1	119	0.43
S8+	118	0.42
B4	117	0.42
Y1	116	0.41
X2.1	113	0.40
C1	110	0.39
N6+	107	0.38
B1	104	0.37
N5+	100	0.36
Q4.3	98	0.35
K1	97	0.35
X3.4	96	0.34
Q1.1	95	0.34
S6+	93	0.33
T1.1.2	90	0.32
X3.2	88	0.31
S1.1.1	85	0.30
Z3	82	0.29
N5.1+	81	0.29
S1.1.2+	80	0.29
A10+	79	0.28
Q1.3	79	0.28
W4	77	0.28
O1.2	75	0.27
H4	72	0.26
X7+	71	0.25
A2.1+	68	0.24
H1	68	0.24
K2	66	0.24
A12+	65	0.23
X5.2+	62	0.22
T3-	62	0.22
O3	62	0.22
A6.1-	61	0.22
O4.3	61	0.22

Semtag	Freq.	Relative Frequency (%)
B5	61	0.22
H2	61	0.22
A6.1+	60	0.21
W1	60	0.21
I3.1	58	0.21
Q4.1	58	0.21
E2++	57	0.20
N5++	54	0.19
A9-	53	0.19
T1	53	0.19
X2.5+	53	0.19
H5	53	0.19
T2-	52	0.19
I2.2	50	0.18
Z7	50	0.18
O4.6+	50	0.18
X8+	48	0.17
T3	47	0.17
O1.1	46	0.16
T1.3-	44	0.16
L2	44	0.16
N3.2-	44	0.16
N6	42	0.15
X4.1	42	0.15
A11.1+	42	0.15
A15-	41	0.15
E4.1+	40	0.14
B3	38	0.14
O4.2+	37	0.13
A13.2	37	0.13
L3	36	0.13
M4	36	0.13
T2+	35	0.13
O4.1	35	0.13
F2	33	0.12
S7.1+	33	0.12
X9.2+	33	0.12
E2+++	32	0.11
M5	32	0.11
N5.1-	31	0.11

Semtag	Freq.	Relative Frequency (%)
A13.4	30	0.11
N3.8+	30	0.11
A5.3+	29	0.10
A12-	29	0.10
O1	28	0.10
B2-	27	0.10
A1.1.2	27	0.10
T1.2	26	0.09
A5.2+	26	0.09
N6+++	26	0.09
K4	25	0.09
O4.4	25	0.09
T3+	25	0.09
T1.1.1	25	0.09
N3.2+	24	0.09
A1.8+	23	0.08
A14	22	0.08
S9	22	0.08
A4.2+	22	0.08
A6.1	21	0.08
A1.2+	21	0.08
N5.2+	21	0.08
G3	21	0.08
A5.1---	21	0.08
K5.2	20	0.07
W5	20	0.07
N3.7+	20	0.07
I1.1	20	0.07
A5.1+++	19	0.07
A5.2-	18	0.06
S1.2.4+	18	0.06
A13.6	18	0.06
N3.3	18	0.06
Q4.2	18	0.06
A9	18	0.06
M8	17	0.06
N6-	17	0.06
I1.3	17	0.06
N3.2	17	0.06
X6+	16	0.06
I1.2	16	0.06
E2-	16	0.06
G1.1	16	0.06

Semtag	Freq.	Relative Frequency (%)
A8	15	0.05
S7.4+	15	0.05
A10-	15	0.05
G2.1	15	0.05
S1.1.3+	14	0.05
O4.6-	14	0.05
N5+++	14	0.05
L1+	13	0.05
X7-	13	0.05
A5.4-	13	0.05
F4	13	0.05
X2.6+	12	0.04
T4+	12	0.04
O4.2-	12	0.04
T2++	12	0.04
A6.3+	12	0.04
L1-	12	0.04
N5-	12	0.04
S3.2	11	0.04
N2	11	0.04
K3	11	0.04
A5.1-	11	0.04
I1	11	0.04
A13.7	11	0.04
N3.2+++	11	0.04
T4-	10	0.04
X2.2-	10	0.04
A6.2-	10	0.04
A1.7+	10	0.04
X4.2	10	0.04
E6-	10	0.04
O4.5	10	0.04
S1.2.5+	10	0.04
E3-	10	0.04
S1.2.1+	9	0.03
E5-	9	0.03
S5-	9	0.03
N5---	9	0.03
X9.2++	9	0.03
K5	8	0.03
A11.2+	8	0.03
T4--	8	0.03
X2.3+	8	0.03
X9.1+	8	0.03
S7.3+	8	0.03
N3.6	8	0.03
N3.1	8	0.03
Q4	7	0.03
O1.3	7	0.03
A12++	7	0.03
W2	7	0.03
A6.1+++	6	0.02
A5.4+	6	0.02
A5.1++	6	0.02
X5.2-	6	0.02
E3+	6	0.02
X3.1+	6	0.02
O1.2-	6	0.02
A5	6	0.02
A13.1	6	0.02
A1.7-	6	0.02
O4.6	6	0.02
N3.7	6	0.02
I2.1	6	0.02
N3.7+++	6	0.02
A15+	6	0.02
T2+++	5	0.02
A1.6	5	0.02
S2.2	5	0.02
X2.6-	5	0.02
A5.3-	5	0.02
B2	5	0.02
E4.1-	5	0.02
S1.1.4+	5	0.02
S7.3	4	0.01
T1.3+	4	0.01
S7.1-	4	0.01
N3.3-	4	0.01
X3.2+	4	0.01
N3.8-	4	0.01
N3.7-	4	0.01
S2.1	4	0.01
X5.1+	4	0.01

Semtag	Freq.	Relative Frequency (%)
N3	4	0.01
G2.1-	4	0.01
X9.2-	4	0.01
S6-	3	0.01
I1.3-	3	0.01
X2	3	0.01
A7	3	0.01
N3.5	3	0.01
T3++	3	0.01
E4.2+	3	0.01
N3.5+	3	0.01
A2.1-	3	0.01
B2+	2	0.01
A6.2--	2	0.01
A13.5	2	0.01
K6	2	0.01
A1.5	2	0.01
S1.2.1-	2	0.01
X5.2++	2	0.01
S1.2.5-	2	0.01
A7+++	2	0.01
S7.2+	2	0.01
X2.5-	2	0.01
H4-	2	0.01
X3.3	2	0.01
T1.3++	2	0.01
N3.4	2	0.01
A15++	2	0.01
H3	2	0.01
G2.2+	2	0.01
N3.3--	2	0.01
G2.2-	2	0.01
A1.4	2	0.01
X1	1	0.00
N3.3---	1	0.00
F1-	1	0.00
S1.2.6-	1	0.00
A11.1++	1	0.00
B2++	1	0.00
N3.8++	1	0.00

Semtag	Freq.	Relative Frequency (%)
I1.1+	1	.00
X3.2-	1	0.00
A13	1	0.00
G2.2	1	0.00
X8+++	1	0.00
A3-	1	0.00
A11.1+++	1	0.00
S1.2.5++	1	0.00
N3.8++++	1	0.00
T1.1	1	0.00
I4	1	0.00
T3---	1	0.00
N3.3+	1	0.00
A1.9	1	0.00
A1.3+	1	0.00
S1.2.4-	1	0.00
T3+++	1	0.00
G1.2	1	0.00
G3-	1	0.00
N3.2++	1	0.00
N3.2--	1	0.00
N5--	1	0.00
A1.5.1+	1	0.00
A5.1	1	0.00
A1.9-	1	0.00
S1.1.3-	1	0.00
I3.1-	1	0.00

The Semtag column in Table 28 shows the abbreviations of the semantic fields. The explanations of these abbreviations have been given in Appendix B. Graphic 9 below shows the frequency of the main semantic fields of the words in the coursebook.
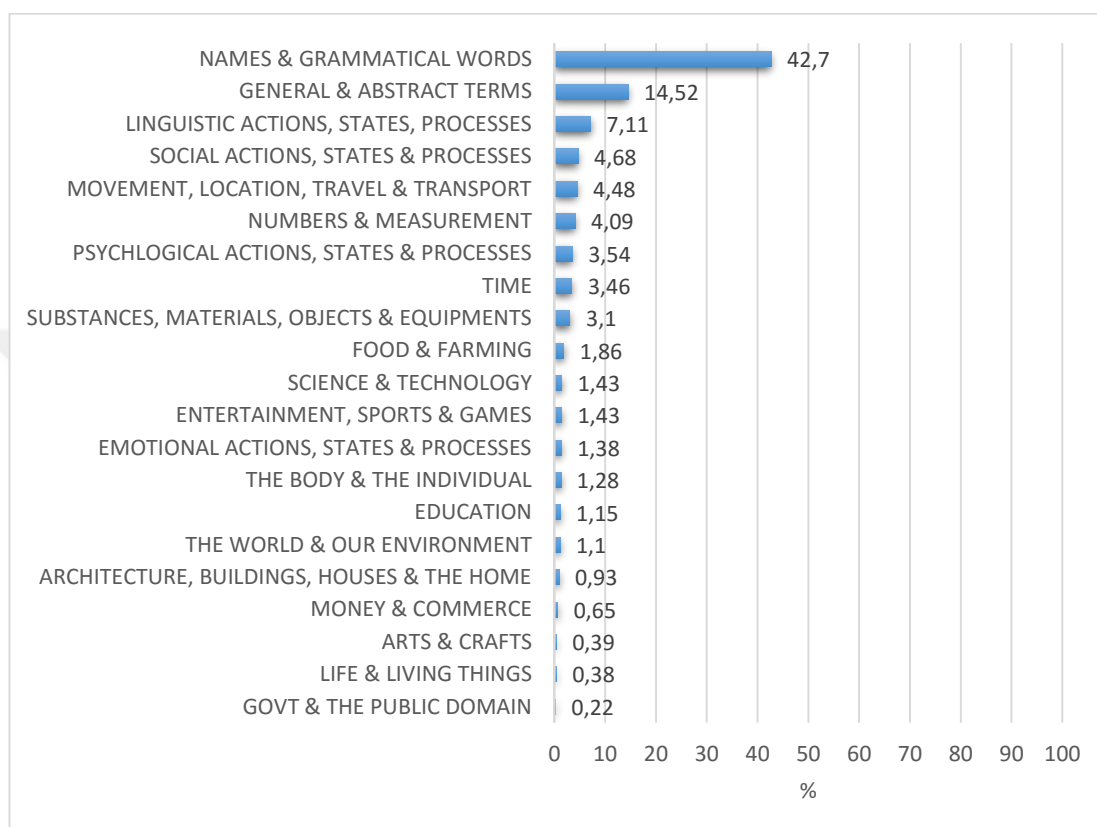
Graphic 9: Frequency of the semantic fields of the words in *Upturn in English 8* in terms of main categories.

Graphic 9 shows that the semantic field "*Names and Grammatical Words*" constitute nearly half of the whole semantic fields with the percentage of 42.7 %. This high percentage is quite natural in that the grammatical words (prepositions, pronouns, articles, determiners, conjunctions, auxiliary verbs and particles) and proper nouns belong to this semantic field and they form the huge percentage of the whole words in any written or spoken corpus.

The semantic field "*Names and Grammatical Words*" is followed by "*General and Abstract Terms*" with the percentage of 14.52% and then comes the semantic field "*Linguistic Actions, States and Processes*" with the percentage of 7.11%. Apart from

these three semantic field categories, the other categories constitute 35.55% of the whole semantic field categories.

It is observed in Graphic 9 that some of the semantic fields are directly reflect the vocabulary of the some of the units in the coursebook. For example, the semantic field "*Social Actions, States and Processes*" reflects Unit 1 (Friendship), the semantic field "*Movement, Location, Travel and Transport*" reflects Unit 7 (Tourism), the semantic field "*Numbers and Measurement*" reflects Unit 9 (Science), the semantic field "*Psychological Actions, States and Processes*" reflects Unit 1 (Friendship) and Unit 2 (Teen Life), the semantic field "*Substances, Materials, Objects and Equipment*" reflects Unit 1 (Friendship), Unit 3 (Cooking) and Unit 9 (Science), the semantic field "*Food and Farming*" reflects Unit 3 (Cooking), the semantic field "*Science and Technology*" reflects Unit 4 (Communication), Unit 5 (Internet) and Unit 9 (Science) and the semantic field "Entertainment, Sports and Games" reflects Unit 6 (Adventures).

Graphic 9 also shows that the least frequent semantic fields according to the vocabulary of the coursebook are the semantic field "*Architecture, Building, Houses and the Home*" with the percentage of 0.93%, the semantic field "*Money and Commerce*" with the percentage of 0.65%, the semantic field "*Life and Living Things*" with the percentage of 0.38% and the semantic filed of "*Government and the Public Domain*" with the percentage of 0.22%. These semantic fields are below the 1% in the frequency list. The low percentage of the semantic fields "*Money and Commerce*", "*Life and Living Things*" and "*Government and the Public Domain*" stems from the fact that there are no units in the coursebook related with these semantic fields. However, although there is a unit in the coursebook named "*Chores*" (Unit 8), the vocabulary of which is related the houses, the home and furniture, the semantic field "*Architecture, Building, Houses and the Home*" has a low percentage of frequency (0.93%). It shows that the vocabulary of the Unit "*Chores*" is not presented successfully enough within the coursebook.

As well as main semantic fields, Table 28 shows the sub-categories of semantic fields. Below is Graphic 10 which shows the most frequently used 20 sub-categories of semantic fields.
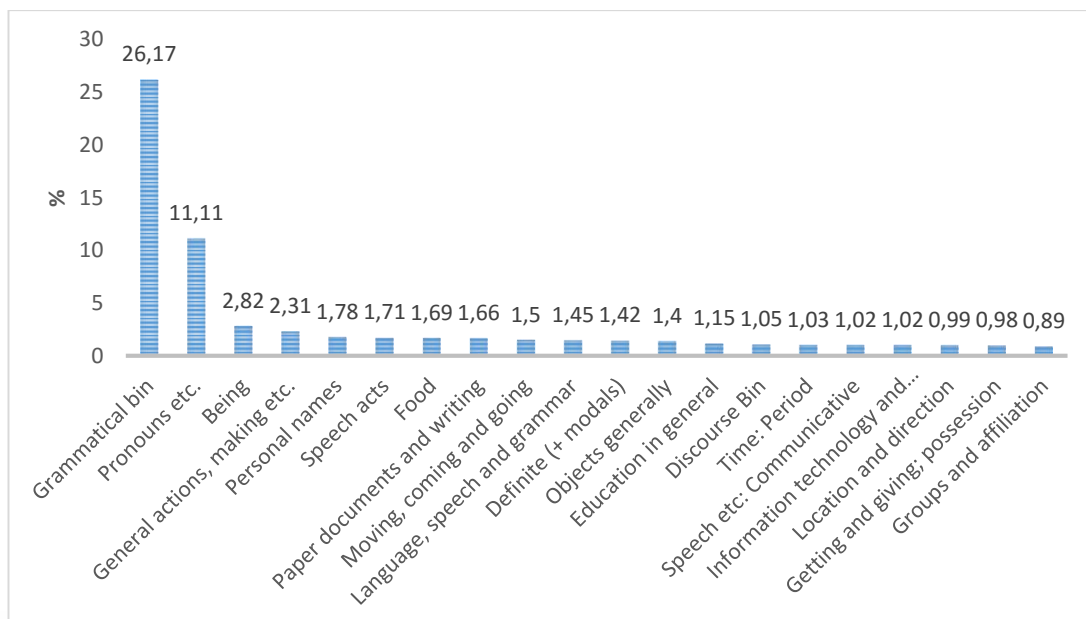
Graphic 10: Most frequent 20 semantic fields of the words in *Upturn in English 8* in terms of sub-categories.

Graphic 10 shows that, as it is in main categories of semantic fields, the sub-categories of the semantic fields show the great domination of grammatical words and proper nouns. The semantic sub-field "*Grammatical bin*" has a percentage of 26.17%. It is consisted of prepositions, articles, determiners, conjunctions, auxiliary verbs, particles and proper nouns. Likewise, the semantic sub-field "*Pronouns*" (e.g. subjective, objective, possessive, reflexive pronouns) has a percentage of 11.11%.

Among the most frequent 20 semantic sub-fields which have been dominated by grammatical words, it is interesting to observe that the semantic sub-category "*Food*" is in the 7th rank with the percentage of 1.69%. It shows that the vocabulary of the Unit "*Cooking*" has been well-presented in the coursebook. There are 474 tokens in the coursebook related to the semantic field "*Food*". Moreover, the semantic sub-category "*Moving, Coming and Going*" is in the 9th rank with the percentage of 1,5%. There are 420 tokens in the coursebook related to this filed. These tokens mainly belong to the Unit "*Tourism*" in the coursebook. It can be stated that Unit "Tourism" has been well-presented in terms of vocabulary load. Nevertheless, the vocabulary of the other units in the coursebook *Upturn in English 8* has not been reflected as highly as units "*Cooking*" and "*Tourism*".

In part 4.8, the frequencies of the semantic fields of the vocabulary in the coursebook have been compared to British National Corpus and overused and underused semantic fields compared to BNC have been analysed.

4.6. The comparison of the frequency of the words in the coursebook with BNC sampler

In this part, the corpus of the coursebook *Upturn in English 8* has been compared with the British National Corpus sampler. As it has been stated before, the word, which has the most important relative frequency difference between the coursebook and BNC sampler shows the largest LL value and is placed at the top of the list. Below is Table 29 which shows the comparison of the corpora of the coursebook and BNC sampler in terms of word frequencies of 100 words. The comparison of the whole list has been given in Appendix E.

Table 29: The top 100 most significant differences at word level between the corpora of the coursebook *Upturn in English 8* and BNC Sampler in terms of over usage

Item	Upturn in English 8		BNC Sampler		O/U Use	LL
	Freq.	Rel. Freq.(%)	Freq.	Rel. Freq. (%)		
unit	194	0.69	38	0.00	+	1187.02
internet	99	0.35	0	0.00	+	710.24
write	136	0.49	226	0.02	+	509.15
sentences	63	0.23	2	0.00	+	434.22
your	348	1.24	3250	0.33	+	392.05
activity	68	0.24	29	0.00	+	371.13
text	63	0.23	20	0.00	+	361.43
the	1749	6.25	37283	3.79	+	359.71
photos	52	0.19	2	0.00	+	356.06
below	69	0.25	42	0.00	+	350.13
lesson	60	0.21	30	0.00	+	317.56
read	94	0.34	229	0.02	+	297.65
science	52	0.19	19	0.00	+	291.64
questions	76	0.27	144	0.01	+	269.70
am	94	0.34	281	0.03	+	267.85
classmates	36	0.13	0	0.00	+	258.27
check	66	0.24	106	0.01	+	250.39

chores	37	0.13	3	0.00	+	244.30
simple	55	0.20	59	0.01	+	239.99
example	52	0.19	49	0.00	+	235.88
favorite	32	0.11	0	0.00	+	229.57
movies	33	0.12	1	0.00	+	227.78
share	57	0.20	87	0.01	+	220.48
sports	41	0.15	20	0.00	+	218.08
online	35	0.13	8	0.00	+	210.23
usually	65	0.23	171	0.02	+	198.11
partner	37	0.13	18	0.00	+	196.91
use	107	0.38	614	0.06	+	196.56
my	215	0.77	2354	0.24	+	196.44
prefer	41	0.15	38	0.00	+	186.87
ok	37	0.13	23	0.00	+	186.85
friends	54	0.19	111	0.01	+	185.00
earthquake	27	0.10	2	0.00	+	179.26
communication	31	0.11	10	0.00	+	177.41
answer	58	0.21	156	0.02	+	174.79
checklist	24	0.09	0	0.00	+	172.18
tourism	25	0.09	1	0.00	+	170.93
student	35	0.13	28	0.00	+	166.11
information	67	0.24	262	0.03	+	162.81
city	27	0.10	6	0.00	+	162.75
teen	22	0.08	0	0.00	+	157.83
p.	22	0.08	0	0.00	+	157.83
activities	37	0.13	43	0.00	+	157.40
short	43	0.15	76	0.01	+	157.06
match	35	0.13	37	0.00	+	153.41
invitation	27	0.10	10	0.00	+	151.08
are	306	1.09	4812	0.49	+	148.14
Dan	24	0.09	5	0.00	+	145.80
follow	38	0.14	60	0.01	+	145.11
email	20	0.07	0	0.00	+	143.48
adventures	20	0.07	0	0.00	+	143.48
natural	32	0.11	32	0.00	+	142.65
study	32	0.11	34	0.00	+	140.05
can	254	0.91	3758	0.38	+	139.70
paragraph	34	0.12	47	0.00	+	136.37
modem	19	0.07	0	0.00	+	136.31
about	230	0.82	3284	0.33	+	135.65
movie	23	0.08	7	0.00	+	132.80
texts	18	0.06	0	0.00	+	129.13
rafting	18	0.06	0	0.00	+	129.13
noun	18	0.06	0	0.00	+	129.13
dishes	29	0.10	30	0.00	+	127.96
friendship	23	0.08	9	0.00	+	127.49

boxes	29	0.10	31	0.00	+	126.68
school	64	0.23	348	0.04	+	122.83
listen	44	0.16	143	0.01	+	119.64
daily	21	0.08	7	0.00	+	119.56
is	490	1.75	9985	1.02	+	118.07
teens	18	0.06	2	0.00	+	116.24
hi	28	0.10	36	0.00	+	115.18
make	83	0.30	645	0.07	+	115.05
messages	23	0.08	15	0.00	+	114.87
extreme sports	16	0.06	0	0.00	+	114.79
bike	31	0.11	53	0.01	+	114.75
chart	18	0.06	3	0.00	+	112.08
brochure	18	0.06	3	0.00	+	112.08
forces	19	0.07	5	0.00	+	112.03
cycling	17	0.06	2	0.00	+	109.29
mom	15	0.05	0	0.00	+	107.61
tv	16	0.06	1	0.00	+	107.24
class	29	0.10	51	0.01	+	106.14
help	53	0.19	282	0.03	+	103.48
ingredients	17	0.06	4	0.00	+	101.73
tourist attractions	14	0.05	0	0.00	+	100.44
parkour	14	0.05	0	0.00	+	100.44
dossier	14	0.05	0	0.00	+	100.44
understand	45	0.16	200	0.02	+	100.37
connection	20	0.07	13	0.00	+	99.96
pairs	18	0.06	7	0.00	+	99.88
with	239	0.85	4031	0.41	+	98.52
phrases	16	0.06	3	0.00	+	98.38
Sam	16	0.06	3	0.00	+	98.38
complete	29	0.10	61	0.01	+	98.34
in	489	1.75	10563	1.07	+	95.71
statements	20	0.07	16	0.00	+	94.92
city	27	0.10	54	0.01	+	93.62
recording	22	0.08	26	0.00	+	93.08
hello	41	0.15	180	0.02	+	92.23
world	42	0.15	196	0.02	+	90.50
box	33	0.12	109	0.01	+	88.89

As it has been stated before, the words above whose frequencies are unusually high in comparison with British National Corpus are called keywords. These words are not the most frequent words, rather, they are unusually frequent.

When Table 29 is examined, it is seen that all of the words above are overused when compared to BNC sampler. One of the reasons of this over usage of the words is because of the fact that some of these words are specific words of the units in the coursebook. For example, the words *activity, photos, tourism, information, city, activities, brochure, tourist* and *world* belong to Unit "Tourism", the words *unit, internet, text, check, share, communication, follow, email, box, modem, messages, TV, chart, online* and *connection* belong to units "Internet" and "Communication", the words *adventure, rafting, extreme, sports, bike, cycling, parkour* and *sports* belong to Unit "Adventures", the words *friendship, teen* and *movies* belong to the Unit "Friendship", the words *favorite, prefer, ingredients* and *dishes* belong to Unit "Cooking", the words *natural* and *earthquake* belong to Unit "Natural Forces", the words *daily* and *chores* belong to Unit "Chores" and the word *science* belongs to Unit "Science". As it is seen, there is an inconsistency in the words that are repeated unusually high. While the units "Internet" and "Communication" have 14 keywords, Unit "Tourism" has 9 keywords and Unit "Adventures" has 8 keywords, the other units do not have enough keywords in the top 100 most significant differences at word level. The high repetitiveness of the keywords helps the vocabulary learning of the students. Nevertheless, there should be unity in the number of the keywords for each unit. There needs to be the same number of keywords for each unit within the coursebook.

When Table 29 is analysed further, it is seen that the words *unit, write, sentences, below, lesson, read, questions, classmates, example, partner, use, match, study* and *listen* are on the top of the list. These words' being on the top of the list results from their usage in the instructions. These words are generally used in the instructions of the activities in the coursebook, thus, their being repetitive may not be on purpose.

Table 29 also shows that, the words which are always used in the spoken language like interjections and discourse markers are very few in the list. The words *ok*, *hi*, *mom* and *hello* have been used in higher frequency when compared to the BNC sampler. However, there are only four words in the 100 words list. This shows the lack of keywords which is necessary for the communicative purposes in language teaching.

Another interesting result is in the usage of definite article "*the*". While the relative frequency of the definite article "*the*" is 6.25% in the coursebook *Upturn in English 8*, the relative frequency of it is 3.79% in the BNC sampler. The results show the high

over usage of definiteness in the vocabulary of the coursebook when compared to the British National Corpus sampler.

Apart from the overused vocabulary in the coursebook when compared to the BNC sampler, there is also underused vocabulary, which has been used unusually low. Below is Table 30, which shows the most underused 50 words in the coursebook with respect to the BNC sampler.

Table 30: The top 50 most significant differences at word level between the corpora of the coursebook *Upturn in English 8* and BNC Sampler in terms of under usage

Item	Upturn in English 8		BNC Sampler		O/U Use	LL
	Freq.	Rel. Freq.(%)	Freq.	Rel. Freq. (%)		
's	3	0.01	20857	2.12	-	1133.33
that	80	0.29	21117	2.15	-	706.99
yeah	5	0.02	9494	0.97	-	483.36
it	347	1.24	26228	2.67	-	261.54
oh	6	0.02	5533	0.56	-	259.73
got	2	0.01	4965	0.51	-	257.81
was	38	0.14	8052	0.82	-	241.41
we	96	0.34	11364	1.16	-	217.29
no	19	0.07	5925	0.60	-	212.66
I	514	1.84	31907	3.25	-	198.55
know	6	0.02	3985	0.41	-	176.77
right	4	0.01	3298	0.34	-	152.12
just	15	0.05	4076	0.41	-	138.24
they	108	0.39	9778	1.00	-	133.30
so	42	0.15	5570	0.57	-	119.14
get	14	0.05	3458	0.35	-	112.25
he	87	0.31	7890	0.80	-	107.84
well	48	0.17	5626	0.57	-	106.45
yes	28	0.10	4297	0.44	-	104.05
said	8	0.03	2758	0.28	-	102.72
did	22	0.08	3841	0.39	-	102.19
had	6	0.02	2554	0.26	-	101.77
but	68	0.24	6561	0.67	-	98.03
were	11	0.04	2867	0.29	-	95.44
one	43	0.15	5026	0.51	-	94.81
there	71	0.25	6504	0.66	-	90.24
be	58	0.21	5700	0.58	-	87.32
this	59	0.21	5416	0.55	-	75.37

if	50	0.18	4612	0.47	-	64.64
think	31	0.11	3491	0.36	-	63.22
all	35	0.13	3689	0.38	-	61.82
you know	9	0.03	1975	0.20	-	60.36
actually	2	0.01	1228	0.12	-	53.60
want	9	0.03	1788	0.18	-	51.65
now	28	0.10	2997	0.30	-	51.17
she	68	0.24	5074	0.52	-	49.30
those	1	0.00	1025	0.10	-	48.85
says	1	0.00	977	0.10	-	46.25
only	4	0.01	1191	0.12	-	41.97
done	2	0.01	978	0.10	-	40.48
thought	2	0.01	935	0.10	-	38.24
quite	2	0.01	928	0.09	-	37.88
which	19	0.07	2081	0.21	-	36.51
them	46	0.16	3463	0.35	-	34.26
here	12	0.04	1576	0.16	-	33.41
still	2	0.01	817	0.08	-	32.16
up	3	0.01	885	0.09	-	31.07
could	21	0.08	2030	0.21	-	30.41
then	50	0.18	3519	0.36	-	30.16
money	1	0.00	643	0.07	-	28.34

When Table 30 is examined, it is seen that the most 50 underused words of the coursebook with respect to BNC are not related with the core vocabulary of the units. The under used vocabulary is generally consisted of articles, determiners, prepositions, past forms of the general verbs, pronouns and conjunctions. It shows that the core vocabulary of the ten units in the coursebook is generally presented as over-usage, that is to say, the words have been presented as key words. Nevertheless, as it has been stated before, there is an inconsistency in the distribution of the keywords within the coursebook. While some of the units have more keywords in the coursebook, some units have very low percentage of key words.

Table 30 shows that the *possessive* "*s*" is the most underused vocabulary of the coursebook when compared with the BNC sampler. It has been repeated only three times in the book with the relative frequency of 0.01%. Although possessive pronouns are used high in number in the coursebook, the *possessive* "*s*" has not been preferred to show possessiveness in the vocabulary of the coursebook. The possessive "s" is followed by singular determiner "that" with the relative frequency of 0.29% and then

comes interjection "yeah" with the relative frequency of 0.02%. The comparison of the parts of speech with BNC sampler has been examined in the section 4.7.

Apart from the comparison of the words in terms of over and underuse, there are two interesting results about the use of the seven days of the week and occurrences of color terms in the coursebook with respect to the BNC sampler.
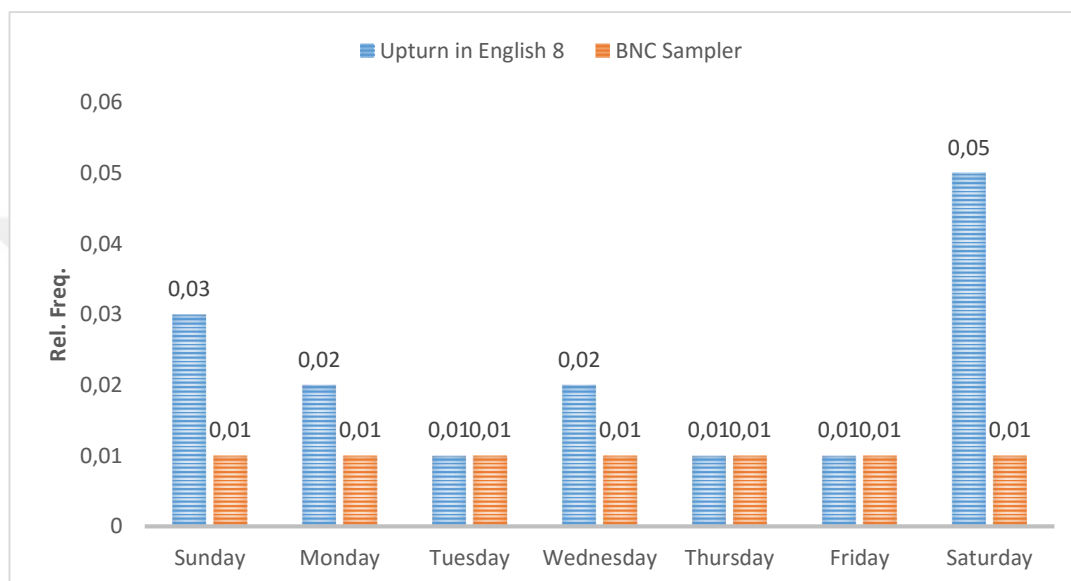
Graphic 11: The use of the seven days of the week in *Upturn in English 8* and BNC sampler

According to Graphic 11, *Saturday* is the most frequently used day of the week in the coursebook and is used unusually frequent when compared with the BNC sampler. It is followed by *Sunday*, *Monday* and *Wednesday*. Although, the relative frequencies of the days of the week in the BNC sampler are the same with the relative frequency of 0,01%, when the frequency of the words within the BNC sampler is examined, it is seen that *Friday* is the most frequent day of the week. It is followed by *Saturday*, *Monday* and *Sunday*. O'Keeffe, McCarthy and Carter (2007) state that there may be cultural reasons for the inconsistent distributions of the days within a corpus. For example, in westernised societies, *Saturday* and *Sunday* are considered as the end of the week and these days are associated with leisure. Likewise, *Monday* is considered as the start of the week for working. Thus, it is seen that the vocabulary of the coursebook reflects the culture of the target language at least in terms of frequency of

the days of the weeks. This result shows that corpus may work as a cultural window in language teaching.

Another interesting result is in the occurrences of color terms in the coursebook with respect to the BNC sampler.
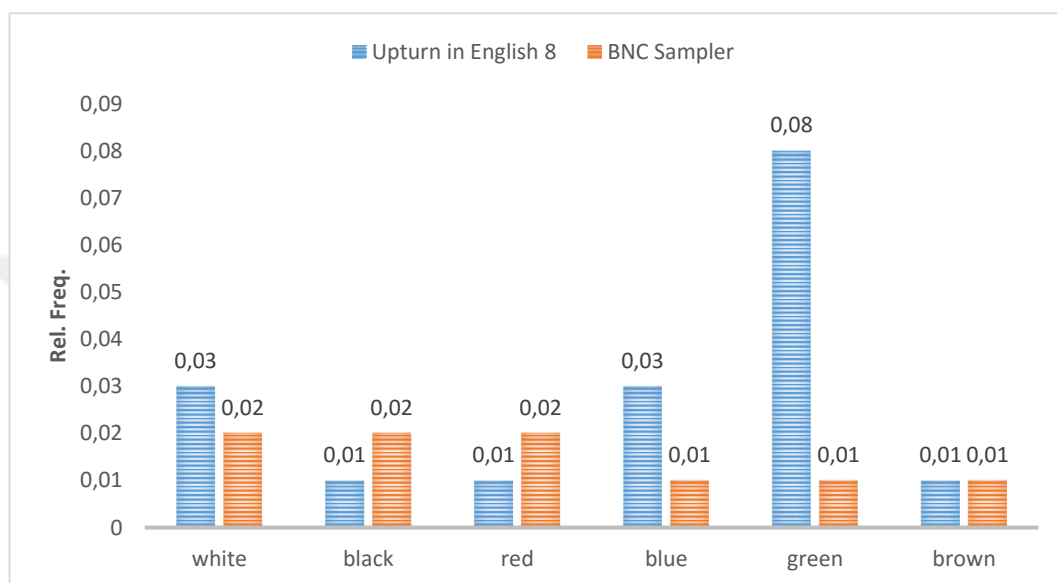
Graphic 12: The usage of the color terms in *Upturn in English 8* and the BNC sampler

Graphic 12 shows that *green* is the most frequent color in the coursebook with relative frequency of 0.08% and it is unusually frequent when compared to the BNC sampler. It is followed by *white* and *blue*. Contrary, when the frequency of the words is examined in the BNC sampler, it is seen that *red* is the most frequent colour. It is followed by *black* and *white*. It is clear that the color terms are not distributed randomly in any corpus. O'Keeffe, McCarthy and Carter (2007) state that the most frequently used colours in Britain also appear in the daily expressions or metaphorical language like *white/black coffee*. The highly occurrence of the *green* in the coursebook may be because of its being typical characteristic of the physical surroundings of Turkey. It is clear that cultural differences play an important role in the distribution of the color items within a corpus.

4.7. The comparison of the frequency of the parts of speech in the coursebook with the BNC sampler

In section 4.4., the frequencies of the parts of speech have been analysed in terms of 132 parts of speech categories. In this section, the over use and under use of the parts of speech of the vocabulary in the coursebook with respect to BNC sampler have been analysed. The whole list of the comparison of the parts of speech between *Upturn in English 8* and BNC sampler has been given in Appendix F. Below is Table 31 which shows the 10 most overused parts of speech in the coursebook when compared to BNC sampler.

Table 31: The top 10 most significant differences at *parts of speech* level of the words between the corpora of the coursebook *Upturn in English 8* and the BNC Sampler in terms of over usage

		Upturn in English 8		BNC Sampler			
	Item	Freq.	Rel. Freq.(%)	Freq.	Rel. Freq.(%)	O/U-use	LL
1	Singular Common Noun	5188	18.55	84446	8.59	+	2325.21
2	Plural Common Noun	1991	7.12	24407	2.48	+	1533.92
3	Base Form Of Lexical Verb	1811	6.47	26118	2.66	+	1047.66
4	Possessive Pronoun	737	2.63	10384	1.06	+	445.81
5	Article	1762	6.30	38652	3.93	+	324.96
6	General Adjective	1554	5.55	36947	3.76	+	202.16
7	General Preposition	1570	5.61	40428	4.11	+	132.67
8	Unit Of Measurement	43	0.15	159	0.02	+	108.25
9	Singular Proper Noun	666	2.38	15229	1.55	+	103.49
10	With, Without	241	0.86	4230	0.43	+	89.95

Table 31 shows that *singular common noun* is the part of speech which has been used most unusually high in the corpus relative to the BNC sampler. It is followed by *plural common noun* and *base form of lexical verb*. These unusually high occurrences of three parts of speech may result from the words are the core vocabulary items of the units, that is to say, they are the key words of the units. Thus, this unusual high occurrence of these top three parts of speech is quite normal.

Apart from these parts of speech, it is seen from Table 31 that *possessive pronoun* is unusually over used when compared to the BNC. It has been stated in part 4.4. that the usage of possessive pronoun, pre-nominal (e.g. my, your, our) is very high in the coursebook. The possessive pronouns, prenominal have totally been used 737 times with the relative frequency of 2.63%. In this part, the use of possessive pronouns has been compared to BNC sampler and it has been found that apart from first person plural possessive pronoun (our), the other possessive pronouns have been over used when compared to BNC sampler (See Graphic 13)
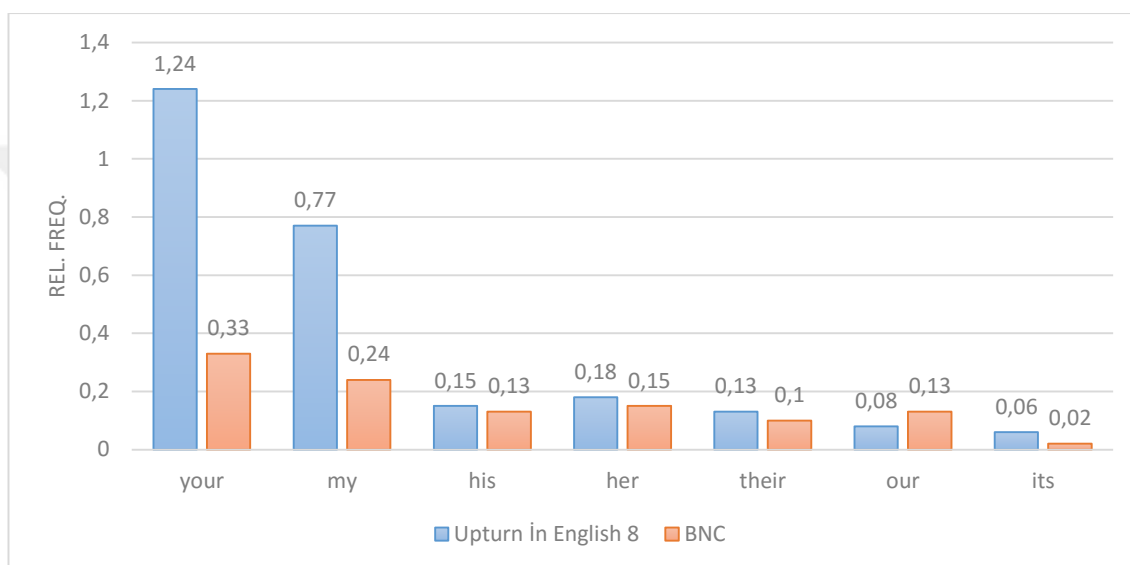
Graphic 13: The comparison of possessive pronouns in *Upturn in English 8* with BNC sampler

Graphic 13 shows that while the relative frequency of second person possessive pronoun "*your*" is 1,24% in the coursebook, it is 0,33% in the BNC sampler. Likewise, the relative frequency of first person singular possessive pronoun "*my*" is 0.77% in the coursebook while the relative frequency of "*my*" is 0.24% in the BNC sampler. Apart from these two possessive pronouns, the other possessive pronouns show similar frequencies within two corpora.

Apart from possessive pronoun, it is seen from Table 31 that articles are also used unusually high in the coursebook. In section 4.4, it has been stated that the definite article "*the*" has been used more than (two times higher than) the indefinite articles "*a*" and "*an*" in the coursebook. The similar results have been detected when the use of articles in the coursebook has been compared to BNC sampler (See Graphic 14).
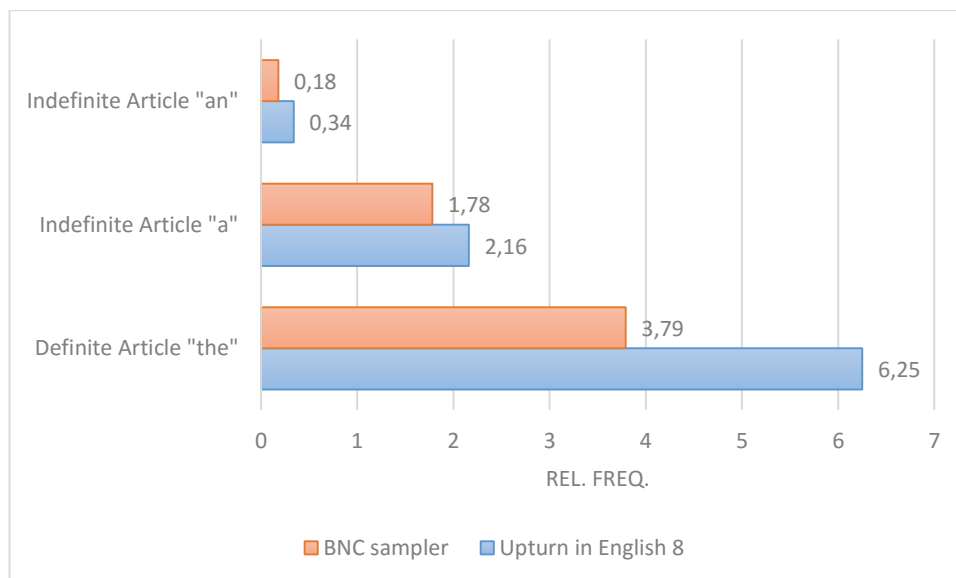
Graphic 14: The comparison of definite article" *the*" and indefinite articles "*a*" and "*an*" in the coursebook with the BNC sampler

It is seen from Graphic 14 that the definite article "*the*" has been used too often in the coursebook with relative frequency of 6,25%. The relative frequency of it is 3,79% in the BNC sampler. Additionally, the indefinite articles "*a*" and "*an*" are also used more highly in the coursebook than in the BNC sampler but not as much as definite article is used. This result shows the unusually high occurrence of the definite article "*the*" in the vocabulary in the coursebook.

As well as over-use of the parts of speech in the coursebook, there are parts of speech which have been underused in the coursebook when compared to the BNC sampler. Table 32 below shows the 10 most underused parts of speech in the coursebook relative to the BNC sampler.

Table 32: The top 10 most significant differences at *parts of speech level* between the corpora of the coursebook *Upturn in English 8* and the BNC Sampler in terms of under use

		Upturn in English 8		BNC Sampler			
	Item	Freq.	Rel. Freq.(%)	Freq.	Rel. Freq.(%)	O/U-use	LL
1	Interjection	140	0.50	33336	3.39	-	1062.87
2	Singular Determiner	109	0.39	21326	2.17	-	610.40
3	Not, N't	135	0.48	18780	1.91	-	419.31
4	General Adverb	598	2.14	39925	4.06	-	301.98

5	Personal Pronoun-It	342	1.22	26247	2.67	-	269.70
6	Past Participle Of Lexical Verb	169	0.60	16468	1.68	-	249.40
7	Was	38	0.14	8118	0.83	-	244.49
8	Personal Pronoun-We	96	0.34	11292	1.15	-	214.46
9	Cardinal Number	108	0.39	11888	1.21	-	209.76
10	Personal Pronoun-I	529	1.89	32828	3.34	-	204.10

Table 32 gives critical information about the under use of *Interjections* in the coursebook *Upturn in English 8*. As it is seen from Table 32 that the Log-Likelihood value of the Interjection is 1062.87 and it is on the top of the list. This means that the part of speech *Interjection* is highly underused relative to the British National Corpus sampler. Interjections are the indispensable part of spoken language and foster communication in the target language. This very low use of *Interjections* in the vocabulary of the coursebook has a negative effect on the communicative purposes of the language teaching. Below is Graphic 15 which shows the whole interjections that are used in the coursebook.
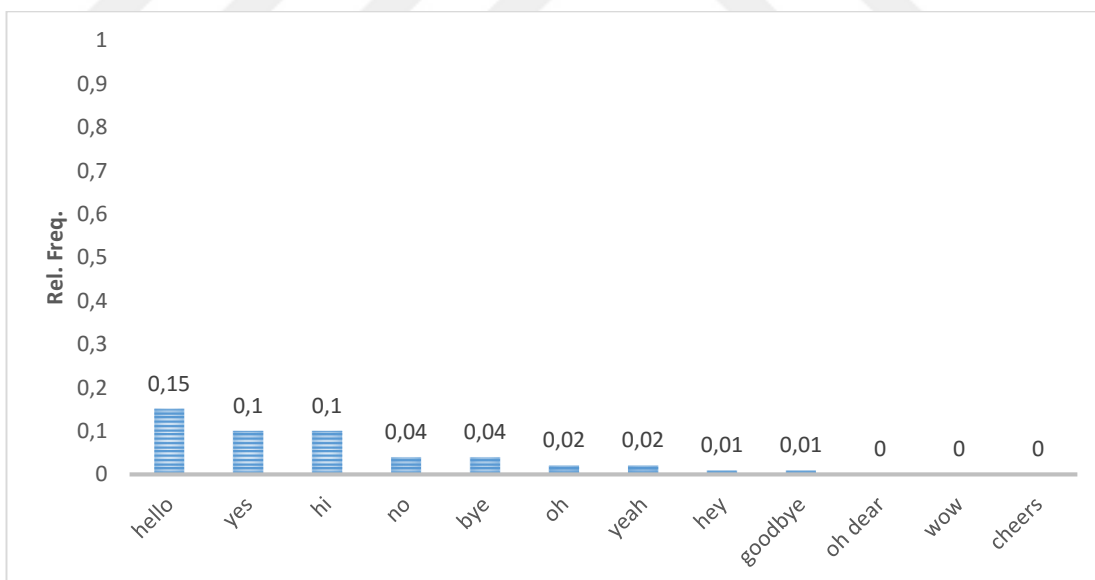
Graphic 15: Interjections used in the coursebook *Upturn in English 8*

Graphic 15 shows the interjections used in the coursebook. As it has been stated in Section 4.6, some of the interjections are highly used when compared to the BNC sampler, however they are very low in number. For example, "*Hello*" is used 41 times, "*Yes*" is used 28 times, "*Hi*" is used 28 times and "*No*" is used 19 times in the

coursebook. The other interjections are used less than ten times in the coursebook. It is necessary that the usage of interjections in the coursebook should be revised for effective language teaching and learning.

Table 32 also shows the under use of singular determiners in the coursebook. There are four singular determiners used in the coursebook. These determiners are *another*, *each*, *this* and *that*. Among these singular determiners, only the determiner "*each*" has been overused when compared to the BNC sampler. However, this over use is too slight in percentage as is seen in Graphic 16.
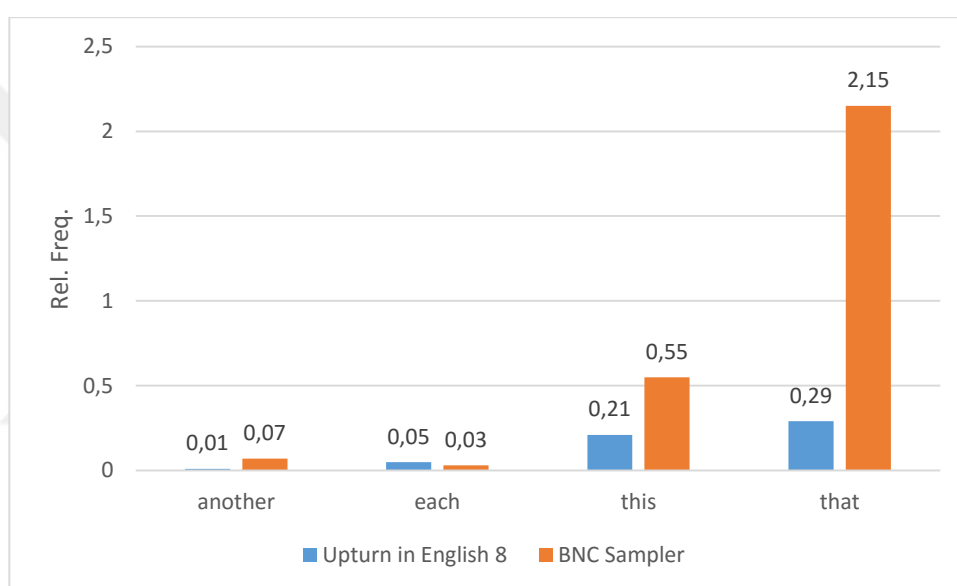
Graphic 16: The use of singular determiner in *Upturn in English 8* and the BNC sampler

As it is seen from Graphic 16, "*that*" is unusually underused in the coursebook with relative frequency of 0,29% though the relative frequency of its usage is 2,15% in the BNC sampler. Likewise, "*this*" and "*another*" are also underused in the coursebook.

When Table 32 is examined, it is also seen that the personal pronouns are also underused in the coursebook relative to the BNC sampler as is seen in Graphic 17.
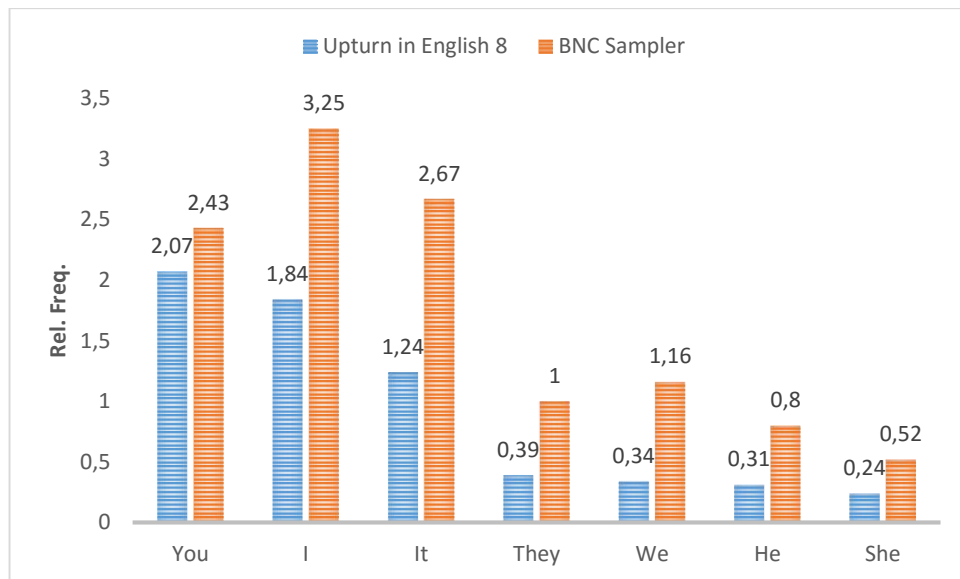
Graphic 17: The use of the personal pronouns in *Upturn in English 8* and the BNC sampler

Graphic 17 shows that the pronouns "*it*", "*we*" and "*I*" are underused relative to BNC sampler. Although the other pronouns (*you*, *they*, *he* and *she*) are also underused when compared to the BNC sampler, it is seen that the percentage of the under use is very slight. It has been stated in Section 4.3 that the pronoun "*I*" is the most frequently used 5th word of the whole corpus in the coursebook. It has been used 514 times in the coursebook with the relative frequency of 1.84%. However, it is seen that it has been unusually underused when compared to the BNC sampler. Although the second person personal pronoun "*you*" is the most frequently used pronoun in the coursebook, the first person singular subjective pronoun "*I*" is the most frequently used pronoun in the BNC sampler.

4.8. The comparison of the frequency of the semantic fields in the coursebook with the BNC sampler

In section 4.5, the frequencies of the semantic fields have been analysed in terms of 21 main and 231 sub-categories of semantic fields. In this section, the over use and under use of the semantic fields of the vocabulary in the coursebook with respect to the BNC sampler has been analysed as well. The whole list of the comparison of the parts of speech between *Upturn in English 8* and the BNC sampler has been given in Appendix G. Below is Table 33 which shows the 10 most overused semantic fields in the coursebook when compared to the BNC sampler.

Table 33: The top 10 most significant differences at semantic field level between the corpora of the coursebook *Upturn in English 8* and the BNC Sampler in terms of over usage

	Item	Upturn in English 8		BNC Sampler		O/U-use	LL
		Freq.	Rel. Freq.(%)	Freq.	Rel. Freq.(%)		
1	Language, speech and grammar	406	1.45	1105	0.11	+	1216.06
2	Information technology and computing	284	1.02	480	0.05	+	1056.13
3	Paper documents and writing	465	1.66	3398	0.35	+	686.14
4	Food	474	1.69	3914	0.40	+	615.73
5	Education in general	323	1.15	1928	0.20	+	574.03
6	Personal relationship: General	174	0.62	752	0.08	+	395.69
7	Speech acts	478	1.71	5794	0.59	+	374.93
8	Objects generally	392	1.40	4156	0.42	+	374.66
9	Sports	169	0.60	914	0.09	+	325.74
10	Science and technology in general	116	0.41	415	0.04	+	298.01

As it has been stated before, the semantic fields of the words which are unusually used too much when compared to a reference corpus are called *key semantic fields*. In Section 4.5, it has been shown that while some of the units have been well presented in terms of vocabulary (e.g. Unit 3 "Cooking", Unit 4 "Communication", Unit 5 "Internet", Unit 7 "Tourism"), some of them (e.g. Unit 8 "Chores") have not been well presented in terms of vocabulary. The same results can be observed when the semantic fields of the words in the coursebook are compared to the BNC sampler.

Table 33 shows that, when the semantic fields which are related to grammar and linguistics have been excluded, the semantic field "*Information Technology and Computing*" has the most Log-Likelihood Value. This semantic field is related with the units "Communication" and "Internet" in the coursebook. It is also seen that the semantic fields "*Food*", "*Sports*" and "*Science and Technology in General*" are the other most unusually frequent semantic fields, in other words, they are the key semantic fields in the coursebook. The semantic field "Food" reflects the Unit "Cooking", "Sports" reflects the Unit "Adventures" and "Science and Technology in

General" reflects the units "Science", "Communication" and "Internet" in the coursebook.

These results show that although the vocabulary of the units "Cooking", "Communication", "Internet", "Science", "Adventures" and "Tourism" are well presented in the coursebook, the vocabulary of the other units are not presented well.

When the semantic fields which have been underused in the coursebook when compared to the BNC are examined in the Appendix G, it is observed that they belong to the semantic filed of "*Grammatical Words*". The comparison of the "*Grammatical Words*" to the BNC has been examined in Section 4.7. Thus, the under usage of the semantic fields relative to BNC has not been given in this section.

In section 4.10, the semantic fields of the vocabulary in the coursebook have been compared to the interests of the 8th graders in order to see whether the semantic fields of the vocabulary in the coursebook correspond to the interests of the students or not.

4.9. The appropriateness of the corpus in the coursebook for A2 Level of CEFR

As it has been stated in Section 2.7.3, Council of Europe (2001: 150-151) presents four criteria to select vocabulary in testing and coursebook materials. These criteria are below:

> 1. To select key words and phrases a) in thematic areas required for the achievement of communicative tasks relevant to learner needs, b) which embody cultural difference and/or significant values and beliefs shared by the social group(s) whose language is being learnt;
> 2. To follow lexico-statistical principles selecting the highest frequency words in large general word-counts or those undertaken for restricted thematic areas;
> 3. To select authentic spoken and written texts and learn/teach whatever words they contain;
> 4. Not to pre-plan vocabulary development, but to allow it to develop organically in response to learner demand when engaged in communicative tasks (Council of Europe, 2001:150-151).

As it is seen, Council of Europe (2001) stresses selecting the words in authentic spoken and written texts in language teaching and learning. That is why, the vocabulary of the coursebook *Upturn in English 8* has been examined by comparing it to the British National Corpus. Likewise, Council of Europe (2001) highlights applying statistical

data in selecting the words with highest frequencies and prioritizes word lists for this purpose. In this study, the most frequently used words in the coursebook have been analysed by referring to wordlists.

Council of Europe (2001:112) presents the vocabulary range of A2 Level of CEFR as follows:

> Has sufficient vocabulary to conduct routine, everyday transactions involving familiar situations and topics, has a sufficient vocabulary for the expression of basic communicative needs and has a sufficient vocabulary for coping with simple survival needs (Council of Europe, 2001:112).

As it is seen, the A2 Level vocabulary range of the CEFR gives great importance to communicative needs and shows that the learner needs to have a vocabulary to deal with the simple survival needs. It has been stated in Section 2.7.3 that vocabulary range of the A2 Level is around 1500 and 2000 words (Van Ek & Trim, 1990; Meara & Milton, 2003; Milton & Alexiou, 2009). Therefore, this vocabulary range needs to be consisted of most useful words for the learners, that is to say, the most frequent words are needed in this context. However, in Section 4.1, it has been shown that there is an inconsistency in the vocabulary profile of the coursebook. Although 76% of the words in the coursebook belong to the class of most frequently used 1000 words in General English, only 9.10% of the words belong to the class of most frequently used second 1000 words in General English. This is a very low percentage because most frequently used second 1000 words are made up of content words and very essential for the meaning.

Although it has been stated in CEFR that the learner needs an adequate vocabulary necessary for basic communication and survival needs, the vocabulary in the coursebook has 12.9% Off-list words. Uncommon vocabulary, words belonging to different dialects and special vocabulary do not help the vocabulary learning of the learners because dealing with Off-List Words is time consuming and not meaningful.

The A2 Level of CEFR further claims that vocabulary is very fundamental to achieve communicative objectives of the language teaching and for the communicative competency of the students. (Council of Europe,2001). To achieve these communicative objectives, the learner needs to take part in the active participation in

communication. Thus, the vocabulary of the spoken language is very necessary. However, as it has been seen in Section 4.7, the interjections, which are the indispensable elements of the conversations, are highly underused relative to the British National Corpus sampler. There are only 12 types of interjection that are used in the coursebook. These interjections are: *hello, yes, hi, no, bye, oh, yeah, hey, goodbye, oh dear, wow* and *cheers*. Bearing this importance of interjections for the communicative purposes of the language teaching, it is necessary that there should be more interjections in the dialogues of the coursebook.

Apart from interjections, it is clearly seen that the vocabulary of the coursebook does not show the examples of the authentic language. As it has been stated earlier, personal pronouns and singular determiners are underused when compared to the British National Corpus. However, the definite and indefinite articles, possessive pronouns and unit of measurements are highly overused relative to the BNC. There seems to be an inconsistency in the usage of the words in the coursebook *Upturn in English 8* with BNC. The vocabulary of the coursebook does not reflect the vocabulary of the BNC.

In conclusion, it can be said that because of the low usage of most frequently used second 1000 words, over use of Off-list words, insufficient use of spoken language (e.g. interjections) and the inconsistency in the use of some parts of speech, the corpus in the coursebook *Upturn in English 8* is not appropriate for A2 level according to CEFR.

4.10. The correspondence of the corpus in the coursebook to the interests of A2 Level English Language Learners.

In this section, the corpus in the coursebook *Upturn in English 8* has been analysed in terms of correspondence to the interests of A2 Level (8th graders) English Language Learners.

Council of Europe (2001:150) states that "constructors of testing and textbook materials are obliged to choose which words to include. Curriculum and syllabus designers are not obliged to do so, but may wish to provide guidelines in the interests of transparency and coherence in educational provision". Bearing this statement in mind, it has been stated in Section 2.8.2 that the new English Language Curriculum

suggests vocabulary for each unit for the coursebook. For example, Ministry of National Education (MEB) (2013) suggests the vocabulary below for Unit 1 "Friendship".

Table 34: Suggested Lexis for Unit 1 in the 8th Grade in ELT Curriculum in Turkey

Unit	Suggested Vocabulary
Unit 1 (Friendship)	back up, best friend, buddy, cool, count on, get on well with somebody, go for a walk, laid-back, mate, pyjama party, secret, sharing, sleepover, slumber party, support

The coursebook writer may get benefit from the suggested vocabulary given by MEB, however, he/she needs to enrich the coursebook with the vocabulary that suits to the interest of the learners. Thus, in this section, the interests of the 527 8th grade English Language Learners from seven regions of Turkey have been analysed. Below is Graphic 18 which shows the interests of the 8th graders for the coursebook.
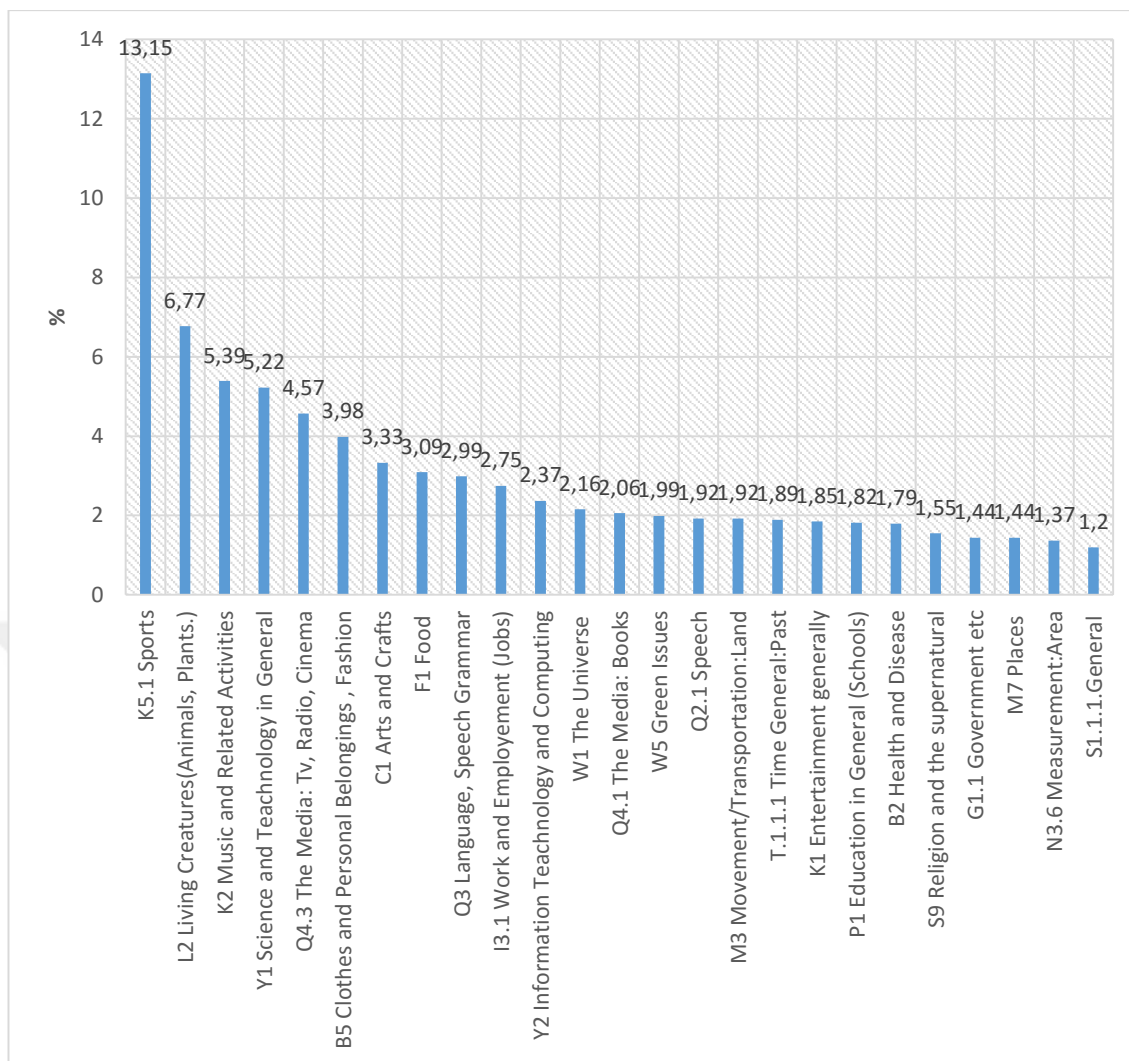
Graphic 18: The interests of A2 Level English Language Learners for the coursebook

It is seen from Graphic 18 that "*Sports*" is the most preferred field that the learners want to see in the coursebook. It has a percentage of 13.15%. It is followed by *Living Creatures (Animals, Plants)* with 6,77%, *Music and Related Activities* with 5,39%, *Science and Technology* with 5.22%, *The Media (TV, Radio, Cinema)* with 4.57%, *Clothes and Fashion* with 3.98%, *Arts and Crafts* with 3,33%, and *Food* with 3.09%.

The list goes on with *Language, Speech, Grammar*, *Work and Job*, *Information Technology and Computing*, *The Universe*, *The Media*: *Books*, *Green Issues*, *Speech*, *Transportation*, *Time: Past*, *Entertainment*, *Education*, *Health*, *Religion and Supernatural*, *Government*, *Places*, *Measurement: Area* and *General Issues*. There are

25 fields in the list, which have the percentage of higher than 1% and make up nearly 85% of the whole fields. There are other fields apart from the ones in the list, however, they are not added to the list because of their low percentages (They have percentages lower than 1%).

In Section 4.5, the frequency of the semantic fields of the vocabulary in the coursebook *Upturn in English 8* have been examined. When the grammatical words and proper nouns have been excluded from the semantic fields of the vocabulary in the coursebook (See Graphic 10), it is seen that the semantic fields *Food*, *Moving, Coming and Going*, *Language, Speech and Grammar*, *Education in General*, *Information Technology and Computing* and *Location and Direction* are the most frequently used semantic fields of the vocabulary in the coursebook.

It is clearly seen that there is inconsistency in the semantic fields of the vocabulary presented in the coursebook and the interests of the 8th graders. Of course it is necessary that the coursebook writers need to refer to the thematic units and the suggested vocabulary given by the Ministry of National Education (MEB) while choosing vocabulary for the coursebooks in foreign language teaching, however, they also need to take the interests of the learners to select the appropriate vocabulary for the units of the coursebook.

Graphic 18 shows that nearly 15% of the students want to see the field *Sports* in the coursebook. Nevertheless, the field *Sport*s is not on the list of 20 most frequent semantic fields of the vocabulary in the coursebook (See Graphic 10). Likewise, *Animals and Plants*, *Music and Related Activities*, *The Media: Tv, Radio, Cinema*, *Clothes and Fashion* and *Arts and Crafts* are the other fields that are most preferred by the eighth graders, but not in the list of 20 most frequent semantic fields of the vocabulary in the coursebook. Only the semantic fields *Food*, *Education in General* and *Information Technology and Computing* reflect the interests of the 8th grade English Language Learners.

When the interests of A2 Level English Language Learners for ELT coursebook are examined in terms of gender, it is seen that female students mostly prefer *Clothes and Personal Belongings*, *Fashion*, *Music and Related Activities*, *The Media, TV, Radio, Cinema*, *Food* and *Health and Disease* while the male students mostly prefer *Sports*,

Living Creatures (Animals, Plants), *Science and Technology in General*, *Work and Employment*, *The Universe*, *Information Technology and Computing*, *Movement/Transportation and Religion* and the *Supernatural*. It is seen that both in terms of male and female analysis, the interest areas do not match with the semantic fields of the vocabulary presented in the coursebook**.**

To sum up, it is seen that the interests of the 8th graders have not been taken into consideration while selecting the vocabulary of the coursebook *Upturn in English 8*.

CHAPTER FIVE

CONCLUSION

5.1. Concluding Remarks

This corpus based study has been conducted to analyse the use of vocabulary in an A2 level EFL coursebook in terms of the use of high frequent words, selection of the words in accordance with the British National Corpus, appropriateness of the words to A2 Level of the CEFR and the correspondence of the vocabulary in the coursebook to the interests of the A2 level English Language Learners. It also aims taking attention of coursebook writers and teachers in selecting vocabulary in foreign language teaching. The vocabulary of the *Upturn in English 8* has been analysed in terms of 10 main categories. These categories are: 1. Vocabulary profile of the whole words, 2. Vocabulary profile of the words in each unit, 3. The frequency of the words in the coursebook, 4. The frequency of the parts of speech in the coursebook, 5. The frequency of the semantic fields in the coursebook, 6. The comparison of the frequency of the words in the coursebook with the BNC sampler, 7. The comparison of the frequency of the parts of speech in the coursebook with the BNC sampler, 8. The comparison of the frequency of semantic fields in the coursebook with the BNC sampler, 9. The appropriateness of the corpus in the coursebook for A2 Level of CEFR and 10. The correspondence of the corpus in the coursebook to the interests of A2 Level English Language Learners.

First of all, the vocabulary profile of the whole words and vocabulary profile of the words in each unit have been analysed. The software programme Compleat Lexical Tutor (Cobb, 2015) has been used to analyse the vocabulary profile. This software programme takes the General Service List formed by Michael West as the basis. The results have shown that there is a clear domination of the K1 Words (0-1000) on the vocabulary of the coursebook. Nearly 76% of the words in the coursebook belong to this class. Thus, it can be claimed that the words in the coursebook correspond with the 1000 most frequent words in General English. However, the words in this class are generally made up of function words (e.g. prepositions, pronouns, determiners,

conjunctions, auxiliary verbs, and particles), thus their over use results from their function, in other words they are not recycled on purpose. Nevertheless, this high percentage disappears when K2 Words (1001-2000) are analysed. K2 Words form only 9.10% of the whole words in the coursebook. The percentage of the K2 Words is very low although they are made up of content words and are more essential for the meaning than K1 Words. It is clearly seen that there should be an increase in the percentage of second 1000 most frequent words. Another interesting result is that Off-List words are unusually used too much with the percentage of 12.09%. There should be a decrease in the percentage of Off-List Words because Off List Words may be proper nouns, uncommon words, abbreviations, acronyms, words belonging to different dialects, special vocabulary, thus, cause problems for vocabulary retention of the learners. Dealing with Off-List Words is time consuming and does not seem so meaningful despite the fact that some of them help the communicative competency of the students, such as the Off-List words in the Unit "Cooking" and Unit "Tourism". Contrary to Off-List Words, AWL (Academic) Words constitute only 3.04%. of the whole words, which is the positive side of the vocabulary profile of the coursebook.

When the vocabulary coverage of each unit is examined separately, it is seen that the percentages of the four categories (K1 Words, K2 Words, Off-List Words and Academic Words) are not balanced. For example, Unit "Friendship" has the most K1 words with the percentage of 82.14%. It is followed by Unit "Natural Forces" with 80.86% and then comes Unit "Teen Life" with 80.75%. However, the units "Science", "Tourism" and "Coking" have very low percentages of K1 Words. In terms of K2 words, Unit "Cooking" has the most K2 Words with the percentage of 13,61%. It is followed by Unit "Adventures" with 11.39% and then Unit "Chores" with 11.22%. Nevertheless, the units "Friendship", "Teen Life", "Communication", "Internet" and "Tourism" have low percentage of K2 Words. Another interesting result is that although Unit "Cooking" has the lowest number of K1 Words; the same unit has the highest number of K2 Words. This imbalance is also seen in the Off-List words and Academic words. Therefore, it can be said that a revision is necessary for the vocabulary choice in terms of most frequent words and distribution of these words in the coursebook. Ünlü (2012) found the same results in her corpus based study about loading, distribution, and repetition patterns of the 2000 high frequency words of general English in 8th grade EFL coursebook. According to her research, there is a

domination of the most frequently used 1000 words on the vocabulary of the coursebook although there is an inconsistency in the distribution of the words in the units. Criado (2009) also has found that there is an inconsistency in the distribution of the words in terms of word lists throughout a Spanish textbook.

After the vocabulary profile in the coursebook has been analysed, *the frequency of the words*, *parts of speech* and *semantic fields* have been examined and then the *frequency of the words*, *parts of speech* and *semantic fields* have been compared to the British National Corpus. The corpus software programme Vmatrix has been used for the frequency analysis and for the comparison to the British National Corpus.

The results of the *frequency of the words* have shown that generally the prepositions, pronouns, determiners, conjunctions, auxiliary verbs and particles are on the top of the list because of their functions. Although the personal pronouns "*I*" and "*you*" are on the top of the list of most frequent 100 words, they do not help to the communicative competence of the students because they are not supported by the discourse markers and interjections in the coursebook.

The results of the *parts of speech* of the words in the coursebook have shown that there is a clear dominance of singular common nouns throughout the words of the coursebook because the most frequently used part of speech in the book is Singular Common Noun (e.g. book, girl) with the relative frequency of 18.55%. It is followed by Plural Common Noun (e.g. books, girls) with the relative frequency 7.12%. Moreover, the results of the *parts of speech* of the words have shown that there is a masculine dominance on the usage of the 3rd person singular subjective personal pronouns because the pronoun "he" has been used 87 times in the coursebook with relative frequency of 0.32%. Contrary, the other pronoun "she" has been used 68 times with the relative frequency of 0.24%. It has also been founded that definiteness plays an important role on the vocabulary of the coursebook because the definite article "the" is used two times higher than the indefinite articles "a" and "and". Apart from the definiteness, the use of possessive pronoun, pre-nominal (e.g. my, your,) is very high in the coursebook. The possessive pronouns, prenominal have been used 737 times with the relative frequency of 2.63%.

The results of the *semantic fields* of the words have shown that while the vocabulary of some of the units in the coursebook have been well presented in the coursebook, the vocabulary of some of the units have not been presented enough. For example, among the most frequent 20 semantic fields of the words in the coursebook, semantic field "Food" is in the 7th rank with the percentage of 1.69%. It shows that the vocabulary of Unit "Cooking" has been well-presented in the coursebook. There are 474 tokens in the coursebook related to the semantic field "Food". Moreover, the semantic field "Moving, Coming and Going" is in the 9th rank with the percentage of 1,5%. There are 420 tokens in the coursebook related to this filed. These tokens mainly belong to the Unit "Tourism" in the coursebook. It can be stated that the Unit "Tourism" has been well-presented in terms of vocabulary load. However, it is not possible to see the same results in all of the units; for example, there is a unit in the coursebook named "Chores" (Unit 8), whose vocabulary items are related to the houses, the home and furniture, nevertheless, the semantic field "Architecture, Building, Houses and the Home" has a low percentage of frequency (0.93%). It shows that the vocabulary of the Unit "Chores" is not presented successfully enough within the coursebook.

The results of the *comparison of the frequency of the words in the coursebook to the BNC* have shown that that the definite article "the" is overused in the coursebook compared to the BNC. While the relative frequency of the definite article "the" is 6.25% in the coursebook *Upturn in English 8*, the relative frequency of it is 3.79% in the BNC sampler. Another result has shown that the possessive "s" is the most underused vocabulary of the coursebook when compared to the BNC. It has been repeated only three times in the book with the relative frequency of 0.01%. Although possessive pronouns are used high in number in the coursebook, the possessive "s" has not been preferred to show possessiveness in the vocabulary of the coursebook.

The results of the *comparison of the parts of speech of the words in the coursebook to the BNC* have shown that apart from first person plural possessive pronoun (our), the other possessive pronouns have been overused when compared to the BNC. Apart from possessive pronouns, definite article "the" has been used unusually high in the coursebook with relative frequency of 6,25%. The relative frequency of it is 3,79% in the BNC sampler. This result shows the unusually high occurrence of the definite article "the" in the vocabulary of the coursebook. In contrast to the over use, there are

underused parts of speech relative to the BNC. For example, part of speech "Interjection" is highly underused relative to the British National Corpus. There are only 12 different interjections (hello, yes, hi, no, bye, oh, yeah, hey, goodbye, oh dear, wow, cheers) used in the coursebook. Likewise, personal pronouns and singular determines are also underused when compared to the BNC.

The *comparison of the semantic fields of the words in the coursebook to BNC* shows the similar results with the already mentioned statement that although the vocabulary of the units "Cooking", "Communication", "Internet", "Science", "Adventures" and "Tourism" is well presented in the coursebook, the vocabulary of the other units is not presented well.

All of these comparisons to the British National Corpus show that the vocabulary of the coursebook *Upturn in English 8* does not reflect the real language samples or authentic usages.

When it comes to the appropriateness of the corpus in the coursebook for A2 Level of CEFR, the results have shown that because of the low use of most frequently used second 1000 words, high use of Off-list words, insufficient use of spoken language (e.g. interjections) and the inconsistency in the usage of some parts of speech, the corpus in the coursebook *Upturn in English 8* is not appropriate for A2 level according to CEFR.

In terms of the correspondence of the corpus in the coursebook to the interests of A2 Level English Language Learners, the results have shown that there is an inconsistency between the semantic fields of the vocabulary presented in the coursebook and the interests of the 8th grade English Language Learners. Nearly 15% of the students want to see the field "Sports" in the coursebook. Nevertheless, the field "Sports" is not on the list of 20 most frequent semantic fields of the vocabulary in the coursebook. Likewise, "Animals and Plants", "Music and Related Activities", "The Media: Tv, Radio, Cinema", "Clothes and Fashion" and "Arts and Crafts" are the other fields that are most preferred by the students, but not on the list of 20 most frequent semantic fields of the vocabulary in the coursebook. Only the semantic fields "Food", "Education in General" and "Information Technology and Computing" reflect the interests of the 8th graders of English.

All in all, the results show that the coursebook *Upturn in English 8* fails to provide 8th grade English Language Learners with vocabulary that reflects the high usage of most frequent words, real and authentic samples of the language. Moreover, it is not appropriate for A2 Level of CEFR and does not take the interests of the learners into account in vocabulary selection.

5.2. Pedagogical Implications of the Study

It has been shown that the vocabulary of the coursebook does not reflect the most frequent second 1000 words in General English. Thus, English Language teachers need to refer to other sources and teaching materials to present the most frequent second 1000 words in teaching.

Apart from teachers, coursebook writers and publishers have a heavy burden on their shoulders. While selecting vocabulary for a coursebook, they need to refer to the wordlists which have the most frequently used words both in spoken and written English. They also need to take into account the criteria of CEFR and the interests of the learners in selecting vocabulary for the coursebook in order to make their learning more meaningful and to motivate them for learning.

5.3. Suggestions for Further Studies

This study has been based on the vocabulary of the 8th grade coursebook. Other studies may analyse the vocabulary profile of the other ELT coursebooks from 5th to 8th grades to make comparisons between the vocabulary profile of the coursebooks. This study also has been based on comparing the frequency of the words, parts of speech and semantic fields to the British National Corpus. Other studies may compare the vocabulary of the ELT coursebooks with different corpora; however, no matter what type of corpora is used, the cultural background in line with learners' language level must be taken into account in the selection of the vocabulary items for the coursebooks.

REFERENCES

Adolphs, S., & Schmitt, N. (2003). Lexical coverage of spoken discourse. *Applied Linguistics* 24(4), 425-438.

Agan, M. (2008). Vocabulary Load Of the English Language Coursebook for Grade 10. Unpublished MA Thesis. Department of English Language Teaching, Uludağ University.

Aijmer, K. (2010). *Corpora and language teaching*. Amsterdam: John Benjamins.

Akbari, R. (2012). Vocabulary in L2 Teaching: Some Statistical Findings. *TESOL Arabia Perspectives*, 19(2), 7-11.

Alcaraz, G. (2009), "Frequency and Functionality: Two Keys for L2 Coursebooks", *International Journal of English Studies*, Special Issue: University of Murcia: 61-72.

Alcaraz-Mármol, G. (2011). Vocabulary input in EFL textbooks: Frequency levels. *Revista Española de LingüísticaAplicada* (RESLA), 21: 147-165. Retrieved from www.um.es/lacell/aelinco/contenido/pdf/52.pdf.

Altıkulaçoğlu, S. (2010). Yabancı dil sınıflarında eşdizimli sözcük öğretimi ve anadilinin rolü. *Dil Dergisi*. 148:37-52.

Arda, A. (2015). *Sunshine 2*. Ankara: Cem Veb Ofset

Arda, A. and Onay, Ö. (2015). *Sunshine 4*. Ankara: Lider Yayıncılık

Arda, A. and Onay, Ö. (2015). *Sunshine 7*. Ankara: Lider Yayıncılık

Arıkan, A. (2009) "Problems with coursebooks in EFL classrooms: Prospective teachers' opinions," *EKEV Academic Review*, 38, 309-315.

Aston, G. and Burnard, L. (1998). *The BNC Handbook: Exploring the British National Corpus with SARA*, Edinburgh University Press, Edinburgh.

Aston, G., Bernardini, S., & Stewart, D. (2004). *Corpora and Language Learners*. Amsterdam: John Benjamins Publishing.

Baleghizadeh, S., & Golbin, M. (2010). The Effect of Vocabulary Size on Reading Comprehension of Iranian EFL Learners. LiBRI. *Linguistic and Literary Broad Research and Innovation*, Vol. 1, Issue 2, 33-46

Baker, P., Hardie, A., & McEnery, T. (2006). *A Glossary of Corpus Linguistics*. Edinburgh: Edinburgh University Press.

Bauer, L. & Nation, P. (1993), Word Families. *International Journal of Lexicography*, Vol. 6 No. 4, 253-279.

Berber Sardinha, T. (1999). Using Key Words in Text Analysis: Practical Aspects. Direct Papers 42, Lael, Universidade Católica de São Paulo

Bernardini, S. (2002), 'Exploring new directions for discovery learning', in Kettemann, B. and Marko, G. Teaching and Learning by Doing Corpus Analysis. Proceedings of the fourth international conference on Teaching and Language Corpora, Graz19–24 July, 2000. Amsterdam: Rodopi, pp. 165–182.

Biber, D. (1993). Representativeness in Corpus Design. *Literary and Linguistic Computing*, 8 (4), Oxford University Press, Oxford, pp. 243 – 257.

Bogaards, P., & Laufer-Dvorkin, B. (2004). *Vocabulary in a second language: Selection, acquisition, and testing*. Amsterdam: John Benjamins Pub.

Bonk, W. (2000). Second language lexical knowledge and listening comprehension. *International Journal of Listening*, *14*,14-31.

Browne, C., Culligan, B., & Phillips, J. (2013). New General Service List. Retrieved from http://www.newgeneralservicelist.org/

Callies, M. (2013). Advancing the Research Agenda of Interlanguage Pragmatics: The Role of Learner Corpora in J. Romero-Trillo (ed.), *Yearbook of Corpus Linguistics and Pragmatics* 2013 (pp. 9-36). Dordrecht: Springer.

Callies, M., & Sandra, G. (2015): "Learner corpora in language testing and assessment: Prospects and challenges". In: Marcus Callies & Sandra Götz (eds), *Learner Corpora in Language Testing and Assessment*. Amsterdam: John Benjamins, 1-9.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book*. New York: American Heritage Publishing Co.

Cephe, P. T., & Toprak, T. E. (2014). The Common European Framework of Reference for Languages: Insights for Language Testing. *Journal of Language and Linguistic Studies*, 10(79-88).

Coady, J., Magoto, J., Hubbard, P., Graney, J., & Mokhtari, K. (1993). High frequency vocabulary and reading proficiency in ESL readers. In T. Huckin, M. Haynes & J. Coady (Eds.), *Second language reading and vocabulary acquisition* (pp.289-298). Norwood, NJ: Ablex.

Cobb, T. (1999). Breadth and depth of vocabulary acquisition with hands-on concordancing. [Electronic version]. *Computer Assisted Language Learning* 12, 345-360. Retrieved November 20, 2004, from http://www.er.uqam.ca/nobel/r21270/cv/Breadth.htm

Cobb, T. (2015). Range for texts v.3 [computer program]. Retrieved from http://www.lextutor.ca/cgi-bin/range/texts/index.pl.

Council of Europe. (2001). Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.

Criado, R. (2009). The Distribution of the Lexical Component in ELT Coursebooks and its Suitability for Vocabulary Acquisition from a Cognitive Perspective. International Journal of English Studies, Special Issue, 39-60.

Crowdy, S. (1993). Spoken corpus design. *Literary and Linguistic Computing* 8 (4): 259-265.

Cubillo, M. C. C., Fortuno, B. B., & Valor, M. L. G. (2010). *Corpus Based Approaches to English Language Teaching*. New York, NY: Continuum.

Cunningham, S., and Moor, P. (2005), *New Cutting Edge*, Harlow: Longman

Cunningsworth, A. (1995). *Choosing your Coursebook*. Oxford: Macmillan Heinemann ELT.

Çakir, İ. (2010). "The Frequency of Culture-Specific Elements in the ELT Coursebooks at Elementary Schools in Turkey. *Novitas-ROYAL (Research on Youth and Language)*, vol.4, pp.182-189

Çelik, S. (2011). Developing Collocational Competence Through Web Based Concordance Activities. *Novitas-Royal (Research on Youth and Language)*, 5(2), 273-286

Dağlıoğlu, Ö. (2015). *İlkokul İngilizce 3*. Ankara: Bilen Görsel Yayıncılık

Dellar, H., & Walkley, A. (2007). *Advanced Coursebook Innovations*. London: Thomson ELT.

Deveci, T. (2004). Why and how to teach collocations. *English Teaching Forum* 42 (2), 16-19

Dilek, M. (2009). An Efl Coursebook Evaluation Study for Vocabulary Teaching in Reading Texts a Case Study: New Bridge to Success, Unpublished MA Thesis. Department of English Language Teaching, Gazi University.

Ellis, R. (1994), *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Engber, C. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing*, 4, 139-155.

Engels, L.K. (1968). The fallacy of word counts. *IRAL* 6: 213-231.

Folse, K. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. Ann Arbor: The University of Michigan Press.

Francis, W. N., & Kucera, H. (1982). *Frequency Analysis of English Usage*. Boston: Houghton Mifflin Company.

Garside, R., Leech, G., and Sampson, G. (eds.) (1987) *The Computational Analysis of English: A Corpus-based Approach*. Longman, London.

Genç, B. (2004). New Trends in Teaching and Learning Vocabulary. *Çukurova University Journal of Social Sciences* 13 (2), 117-126.

Ghadessy, M., Henry, A., & Roseberry, R. L. (2001). *Small corpus studies and ELT: Theory and practice*. Amsterdam: J. Benjamins Pub. Co.

Goulden, R., Nation. P., Nation & Read J. (1990). How large can a receptive vocabulary be? *Applied Linguistics* 11: 341-363.

Granath, S. (2009). Who benefits from learning how to use corpora? In Aijmer, K. (ed.), *Corpora and language teaching*. Amsterdam/Philadelphia: John Benjamins, 47-66.

Granger, S. (1993). International Corpus of Learner English. In Aarts, J., de Haan, P., and Oostdijk, N. (eds.) *English language corpora: Design, analysis and exploitation*. Rodopi, Amsterdam, pp. 57 – 71.

Granger, S. (2002). The contribution of learner corpora to second language acquisition and foreign language teaching. In K. Aijmer (ed.), *Corpora and Language Teaching*, pp.13-32. Amsterdam: John Benjamins.

Grefenstette, G. (1999). Tokenization. In van Halteren, H, (ed.) *Syntactic wordclass tagging*, Kluwer, The Netherlands, pp. 117 – 133.

Günal, O. D., & Demir, C. E. (2012). "Implementation of the New Eighth Grade English Language Curriculum from the Perspectives of Teachers and Students", *Procedia - Social and Behavioral Sciences*, Vol. 47, pp. 1002-1006.

Hana, J., Rosen, A., Štindlová, B., & Jäger, P. (2012). Building a learner corpus. In: Proceedings of the 8th International Conference on Language Resources and Evaluation (LREC 2012). European Language Resources Association.

Harmer, J. (2001). *The practice of English language teaching*. Essex, England: Longman.

Hatch, E., & Brown, C. (1995). *Vocabulary, semantics, and language education*. Cambridge: Cambridge University Press.

Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal*, 36, 153–166. doi: 10.1080/09571730802389983

Hirsh, D., & Nation, P. (1992). What vocabulary size is needed to read unsimplified texts for pleasure? *Reading in a Foreign Language* 8, 2: 689-696.

Hunston, S. (2002). *Corpora in applied linguistics*. Cambridge: Cambridge University Press.

Hymes, D. (1972). On communicative competence. In J. B. Pride & J. Holmes (eds.), *Sociolinguistics*. Harmondsworth: Penguin, 269–293.

İnal, B. (2006), "Coursebook Selection Process and Some of the Most Important Criteria to be Taken into Consideration in Foreign Language Teaching", *Journal of Arts and Sciences*, 5(1), 19-29.

Karatay, H. (2007). Kelime Öğretimi, *Gazi Eğitim Fakültesi Dergisi*, Cilt 27, Sayı: 1, 141-153

Kayaoğlu, M. N. (2011), "A Critical Appraisal of the Language Textbook", *Ahi Evran Üniversitesi Kırşehir Eğitim Fakültesi Dergisi*, 12(4), 341-356.

Kelly, P. (1991). Lexical Ignorance: The Main Obstacle to Listening Comprehension with Advanced Foreign Language Learners. *IRAL*, 29, 135-149.

Kennedy, G. (1998). *An introduction to corpus linguistics*. London: Longman

Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. Journal of Language Teaching and Research, 4,900-913. Retrieved from http://ojs.academypublisher.com/index.php/jltr/article/view/jltr0405900913/755

Kormos, J. (2006). Speech production and second language acquisition. Mahwah, NJ: Lawrence Erlbaum Associates.

Larsen-Freeman, D. (2003). *Techniques and principles in language teaching*. Oxford: Oxford University Press.

Laufer, B. (1986). "Possible Changes in Attitude towards Vocabulary Acquisition", *IRAL*, 24(1), 69-75

Laufer, B. (1989). What percentage of text-lexis is essential for comprehension? In C. Lauren and M. Nordman (eds.), *Special Language: From Humans Thinking to Thinking Machines*. Clevedon: Multilingual Matters.

Laufer, B., &Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16, 307-322.

Leech, G. (2001) The role of frequency in ELT: New corpus evidence brings a re-appraisal. In: ELT in China 2001: Papers presented at the 3rd International Symposium on ELT in China. Foreign Language Teaching and Research Press, Beijing, pp. 1-23.

Leech, G., Rayson, P., and Wilson, A. (2001). *Word frequencies in written and spoken English: based on the British National Corpus*. Longman, London.

Levelt, W. J. M. (1989). *Speaking*. MA: MIT Press.

Little, D. (2006). The Common European Framework of Reference for Languages: content, purpose, origin, reception and impact. *Language Teaching*. 39(3), 167-190.

Little, D. (2011). The Common European Framework of Reference for Languages: A research agenda. Language Teaching, 44, pp 381-393. doi:10.1017/S0261444811000097.

Liu, N., & Nation, P. (1985). Factors affecting guessing vocabulary in context. *RELC Journal*, 16(1), 33-42

McEnery, T., Piao, S., & Xin, X. (2000). Parallel alignment in English and Chinese. In Botley, S., McEnery, A., and Wilson, A. (eds.) *Multilingual corpora in teaching and research*. Rodopi, Amsterdam, pp. 177 – 191.

Meara, P., & Milton, J. (2003). X_Lex, the Swansea Levels Test. Newbury: Express.

MacGregor, A. (2014). Young Learners and Lexical Awareness: Children's Engagement with Wordlists and Concordances. *TESOL Journal*; 5: 120

Martini, J. O. P. (2012). High frequency vocabulary in a secondary Quebec ESL textbook corpus (Unpublished master's thesis). Concordia University, Montreal, Canada

McCarthy, M. (1998). *Spoken Language and Applied Linguistics*. Cambridge: Cambridge University Press

MEB, (2013). *İlköğretim Kurumları (İlkokullar ve Ortaokullar) İngilizce Dersi (2,3,4,5,6,7 ve 8. Sınıflar) Öğretim Programı*. Ankara, Milli Eğitim Basımevi

Michael F. G., Diane, A., & Jeannette, M. (2012). *Teaching Vocabulary to English Language Learners*. New York, NY: Teachers College Press.

Milton, J. (2010) The development of vocabulary breadth across the CEFR levels. A common basis for the elaboration of language syllabuses, curriculum guidelines, examinations, and textbooks across Europe. In Communicative Proficiency and Linguistic Development: Intersections between SLA and Language Testing Research, L. Bartning, M. Martin & I. Vedder (eds) 211–232. Reggio Emila: European Second Language Association.

Milton, J., & Alexiou, T. (2009). Vocabulary size and the Common European Framework of Reference for Languages. In B. Richards et al. (Eds.), *Vocabulary studies in first and second language acquisition* (pp. 194–211). Basingstoke: Palgrave.

Miranda, A. (1990), "The Vocabulary of the English Coursebooks", *Revista Espanola de Linguistica Applicada*, 6, 111-117

Moghadam, H. S., Zainalb, Z., & Ghaderpourc, M. (2012). A Review on the Important Role of Vocabulary Knowledge in Reading Comprehension Performance. The 8th International Language for Specific Purposes (LSP) Seminar - Aligning Theoretical Knowledge with Professional Practice. Procedia - Social and Behavioral Sciences 66: 555 – 563.

Nation, I.S.P. (2000), Learning Vocabulary in Another Language, Victoria University of Wellington: English Language Institute Occasional Publication

Nation, I.S.P. (2006), "How large a Vocabulary is Needed for Reading and Listening", *Canadian Modern Languages Review*, 63(1)59-82.

Nation, I.S.P. (1997), "Vocabulary size, text coverage, and word lists", in Schmitt; McCarthy, *Vocabulary: Description, Acquisition and Pedagogy*, Cambridge: Cambridge University Press

Nazary, M. (2008). The Role of L1 in L2 Acquisition: Attitudes of Iranian University Students. Novitas-ROYAL, 2(2), 138 – 153

Nunan, D. (1991). *Language Teaching Methodology*. New York: Prentice Hall.

O'Keeffe, A., McCarthy, M., and Carter, R. (2007), *From Corpus to Classroom: Language Use and Language Teaching*, Cambridge: Cambridge University Press

Orpin, D. (1997). The lexis of corruption in the news: a corpus-based study in ideology, Unpublished MA Thesis. Department of English, University of Birmingham.

Özbay, A. Ş., & Kayaoğlu, N. (2015). Efl Teacher's Reflections Towards the Use of Computerized Corpora as A Teaching Tool in Their Classrooms. *Cukurova University Faculty of Education Journal*, vol.2015, no.1, pp.2-10

Özturk, M. (2003). "Lexical Competence in the Common European Framework of Reference for Languages". Paper presented at I. International Symposium: Common European Framework and Foreign Language Education in Turkey: Bursa.

Procter, P. (1978). *Longman dictionary of contemporary English*. Harlow [England: Longman.

Rayson, P. (2002). Matrix: A statistical method and software tool for linguistic analysis through corpus comparison, Unpublished Ph.D. thesis. Computing Department, Lancaster University.

Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*. 13:4 pp. 519-549. DOI: 10.1075/ijcl.13.4.06ray

Rayson, P. (2009). Wmatrix: a web-based corpus processing environment, Computing Department, Lancaster University. Retrieved from http://ucrel.lancs.ac.uk/wmatrix/

Richards, J. C. (2001). *Curriculum Development in Language Teaching*. Cambridge: Cambridge University Press.

Römer, U. (2004). Comparing Real and Ideal Language Learner Input: The Use of an EFL Textbook Corpus in Corpus Linguistics and Language Teaching. In G. Aston, S. Bernardini & D. Stewart (eds.). *Corpora and Language Learners*. Amsterdam: John Benjamins. 151-168.

Römer, U. (2010). Using general and specialized corpora in English language teaching: past, present and future. In: Campoy-Cubillo, M. et al. (eds.): *Corpus-based approaches to English Language Teaching*. London: Continuum, 18-38

Sarıca, Ö. (2009). An Analysis of the Language Levels of the Eighth (8th) Graders in State Primary Schools in Turkey According to Common European Framework Criteria and Content, Unpublished MA Thesis. Department of English Language Teaching, Pamukkale University.

Schmitt, N. (2000). *Vocabulary in language teaching*. Cambridge, England: Cambridge University Press

Schonell, F.J., Meddleton, I.G., & Shaw, B.A. (l956). A Study of the Oral Vocabulary of Adults. Brisbane: University of Queensland Press.

Sheehan, A. (2004). Making Sense of Words. *English Teaching Forum*. C. 42. S. 1. s. 1-11.

Shei, C., & Pain, H. (2000), 'An ESL writer's collocational aid', *Computer Assisted Language Learning*, vol. 13, no.2, pp. 167-182.

Sheldon, L. E. (1988). Evaluating ELT Textbooks and Materials. *ELT Journal*, 42, 237-247

Sinclair, J. (1987), Looking Up: An Account of the COBUILD Project in Lexical Computing. London: Collins ELT.

Sinclair, J. (1991). *Corpus, concordance and collocation*. Oxford, England: Oxford University Press.

Sinclair, J., & Renouf, A. (1988). A lexical syllabus for language learning. In R. Carter and M. McCarthy (Eds.) *Vocabulary language teaching*. London: Longman

Soars, L., & Soars, J. (2003). *New Headway English Course*: Oxford: Oxford University Press

Staehr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing, *The Language Learning Journal*, 36(2), 139-152.

Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31*, 577-607.

Şener, M. (2015). *Upturn in English 8*. Ankara: Evrensel İletişim Yayınları

Şilit, G. (2015). *English Net 5*. Ankara: Evrensel İletişim Yayınları

Şilit, G. (2015). *English Net 6*. Ankara: Evrensel İletişim Yayınları

Thornbury, S. (2002), *How to Teach Vocabulary*, Harlow: Longman.

Thorndike, E. L., & Lorge, I. (l944). *The Teacher's Word Book of 30,000 Words. Teachers College*, Columbia University.

Topkaraoğlu, M., & Dilman, H. (2015). The impact of vocabulary instruction on vocabulary size levels of students. *Abant İzzet Baysal Üniversitesi Eğitim Fakültesi Dergisi*, 15 (2), 349-363.

Tribble, C. and Jones, G. (1997). *Concordances in the classroom*. Athelstan, Houston, Texas.

Ünlü, F. (2012). Loading, Distribution, And Repetition Patterns of The 2000 High Frequency Words of General English in an Efl Course Book, Unpublished MA Thesis. Department of English Language and Literature, Karadeniz Technical University.

Van Ek, J. A., & Trim, J. L. M. (1990). *Threshold 1990*, Cambridge: Cambridge University Press

Van Ek, J. A., & Trim, J. L. M. (1990). *Waystage 1990*, Cambridge: Cambridge University Press.

Van Ek, J. A., & Trim, J. L. M. (2001). *Vantage*, Cambridge: Cambridge University Press.

Wallace, C. (2008). Vocabulary: The Key to Teaching English Language Learners to Read. *Education Digest*. 73 (9), 36-39

Wang, S. (2015). An Empirical Study on the Role of Vocabulary Knowledge in EFL Listening Comprehension. *Theory and Practice in Language Studies*, Vol. 5, No. 5, pp. 989-995

Wang, X. (2014). The relationship between lexical diversity and EFL writing proficiency. University of Sydney Papers in *TESOL*, 9, 65-88.

Waring, R., & Nation, P. (1997). Vocabulary size, text coverage, and word lists. In *Vocabulary: Description, Acquisition and Pedagogy* N. Schmitt and M. McCarthy (eds.). Cambridge University Press, Cambridge: 6-19.

West, M. (1953). *A General Service List of English Words*. London: Longman, Green and Co.

White, R. V. (1989). *The ELT curriculum: Design, Innovation and Management*. Cambridge: Blackwell.

Widdowson, H. G. (1990). *Aspects of language teaching*. Oxford: Oxford University Press.

Xue, G. & Nation, I.S.P. (1984). A University Word List. *Language Learning and Communication* 3, (2), 215-229.

Zhang L.J., & Anual S.B. (2008). The role of vocabulary in reading comprehension: the case of secondary school students learning English in Singapore. *RELC Journal*, 39 (1), 51-76.

Zimmerman, C., & Schmitt, N. (2005). Lexical questions to guide the teaching and learning of words. *CATESOL Journal* 17/1: 1-7.

Zorba, M. G. (2012). An Evaluation of Anatolian High Schools' 9th Grade Efl Curriculum and Course Materials Through the Principles of the Common European Framework of Reference for Languages, Unpublished MA Thesis. Department of Foreign Language Teaching, Akdeniz University.

APPENDICES

APPENDIX A

THE EXPLANATIONS OF THE ABBREVIATIONS OF PARTS OF SPEECH

(UCREL CLAWS7 Tagset)

APPGE	possessive pronoun, pre-nominal (e.g. my, your, our)
AT	article (e.g. the, no)
AT1	singular article (e.g. a, an, every)
BCL	before-clause marker (e.g. in order (that), in order (to))
CC	coordinating conjunction (e.g. and, or)
CCB	adversative coordinating conjunction (but)
CS	subordinating conjunction (e.g. if, because, unless, so, for)
CSA	as (as conjunction)
CSN	than (as conjunction)
CST	that (as conjunction)
CSW	whether (as conjunction)
DA	after-determiner or post-determiner capable of pronominal function (e.g. such, former, same)
DA1	singular after-determiner (e.g. little, much)
DA2	plural after-determiner (e.g. few, several, many)
DAR	comparative after-determiner (e.g. more, less, fewer)
DAT	superlative after-determiner (e.g. most, least, fewest)
DB	before determiner or pre-determiner capable of pronominal function (all, half)
DB2	plural before-determiner (both)
DD	determiner (capable of pronominal function) (e.g. any, some)
DD1	singular determiner (e.g. this, that, another)
DD2	plural determiner (these, those)
DDQ	wh-determiner (which, what)
DDQGE	wh-determiner, genitive (whose)
DDQV	wh-ever determiner, (whichever, whatever)
EX	existential there
FO	formula
FU	unclassified word
FW	foreign word
GE	Germanic genitive marker - (' or's)
IF	for (as preposition)
II	general preposition
IO	of (as preposition)
IW	with, without (as prepositions)
JJ	general adjective
JJR	general comparative adjective (e.g. older, better, stronger)
JJT	general superlative adjective (e.g. oldest, best, strongest)
JK	catenative adjective (able in be able to, willing in be willing to)
MC	cardinal number, neutral for number (two, three.)
MC1	singular cardinal number (one)
MC2	plural cardinal number (e.g. sixes, sevens)
MCGE	genitive cardinal number, neutral for number (two's, 100's)

MCMC	hyphenated number (40-50, 1770-1827)
MD	ordinal number (e.g. first, second, next, last)
MF	fraction, neutral for number (e.g. quarters, two-thirds)
ND1	singular noun of direction (e.g. north, southeast)
NN	common noun, neutral for number (e.g. sheep, cod, headquarters)
NN1	singular common noun (e.g. book, girl)
NN2	plural common noun (e.g. books, girls)
NNA	following noun of title (e.g. M.A.)
NNB	preceding noun of title (e.g. Mr., Prof.)
NNL1	singular locative noun (e.g. Island, Street)
NNL2	plural locative noun (e.g. Islands, Streets)
NNO	numeral noun, neutral for number (e.g. dozen, hundred)
NNO2	numeral noun, plural (e.g. hundreds, thousands)
NNT1	temporal noun, singular (e.g. day, week, year)
NNT2	temporal noun, plural (e.g. days, weeks, years)
NNU	unit of measurement, neutral for number (e.g. in, cc)
NNU1	singular unit of measurement (e.g. inch, centimetre)
NNU2	plural unit of measurement (e.g. ins., feet)
NP	proper noun, neutral for number (e.g. IBM, Andes)
NP1	singular proper noun (e.g. London, Jane, Frederick)
NP2	plural proper noun (e.g. Browns, Reagans, Koreas)
NPD1	singular weekday noun (e.g. Sunday)
NPD2	plural weekday noun (e.g. Sundays)
NPM1	singular month noun (e.g. October)
NPM2	plural month noun (e.g. Octobers)
PN	indefinite pronoun, neutral for number (none)
PN1	indefinite pronoun, singular (e.g. anyone, everything, nobody, one)
PNQO	objective wh-pronoun (whom)
PNQS	subjective wh-pronoun (who)
PNQV	wh-ever pronoun (whoever)
PNX1	reflexive indefinite pronoun (oneself)
PPGE	nominal possessive personal pronoun (e.g. mine, yours)
PPH1	3rd person sing. neuter personal pronoun (it)
PPHO1	3rd person sing. objective personal pronoun (him, her)
PPHO2	3rd person plural objective personal pronoun (them)
PPHS1	3rd person sing. subjective personal pronoun (he, she)
PPHS2	3rd person plural subjective personal pronoun (they)
PPIO1	1st person sing. objective personal pronoun (me)
PPIO2	1st person plural objective personal pronoun (us)
PPIS1	1st person sing. subjective personal pronoun (I)
PPIS2	1st person plural subjective personal pronoun (we)
PPX1	singular reflexive personal pronoun (e.g. yourself, itself)
PPX2	plural reflexive personal pronoun (e.g. yourselves, themselves)
PPY	2nd person personal pronoun (you)
RA	adverb, after nominal head (e.g. else, galore)
REX	adverb introducing appositional constructions (namely, e.g.)
RG	degree adverb (very, so, too)
RGQ	wh- degree adverb (how)
RGQV	wh-ever degree adverb (however)
RGR	comparative degree adverb (more, less)

RGT	superlative degree adverb (most, least)
RL	locative adverb (e.g. alongside, forward)
RP	prep. adverb, particle (e.g. about, in)
RPK	prep. adv., catenative (about in be about to)
RR	general adverb
RRQ	wh- general adverb (where, when, why, how)
RRQV	wh-ever general adverb (wherever, whenever)
RRR	comparative general adverb (e.g. better, longer)
RRT	superlative general adverb (e.g. best, longest)
RT	quasi-nominal adverb of time (e.g. now, tomorrow)
TO	infinitive marker (to)
UH	interjection (e.g. oh, yes, um)
VB0	be, base form (finite i.e. imperative, subjunctive)
VBDR	were
VBDZ	was
VBG	being
VBI	be, infinitive (To be or not... It will be ..)
VBM	am
VBN	been
VBR	are
VBZ	is
VD0	do, base form (finite)
VDD	did
VDG	doing
VDI	do, infinitive (I may do... To do...)
VDN	done
VDZ	does
VH0	have, base form (finite)
VHD	had (past tense)
VHG	having
VHI	have, infinitive
VHN	had (past participle)
VHZ	has
VM	modal auxiliary (can, will, would, etc.)
VMK	modal catenative (ought, used)
VV0	base form of lexical verb (e.g. give, work)
VVD	past tense of lexical verb (e.g. gave, worked)
VVG	-ing participle of lexical verb (e.g. giving, working)
VVGK	-ing participle catenative (going in be going to)
VVI	infinitive (e.g. to give... It will work...)
VVN	past participle of lexical verb (e.g. given, worked)
VVNK	past participle catenative (e.g. bound in be bound to)
VVZ	-s form of lexical verb (e.g. gives, works)
XX	not, n't
ZZ1	singular letter of the alphabet (e.g. A,b)
ZZ2	plural letter of the alphabet (e.g. A's, b's)

APPENDIX B

THE EXPLANATIONS OF THE ABBREVIATIONS OF SEMANTIC FIELDS

(USAS Semantic Tagset)

See http://ucrel.lancs.ac.uk/usas/ for more details.

A GENERAL & ABSTRACT TERMS
A1   General
A1.1.1 General actions, making etc.
A1.1.2 Damaging and destroying
A1.2   Suitability
A1.3   Caution
A1.4   Chance, luck
A1.5   Use
A1.5.1 Using
A1.5.2 Usefulness
A1.6   Physical/mental
A1.7   Constraint
A1.8   Inclusion/Exclusion
A1.9   Avoiding
A2   Affect
A2.1   Affect: Modify, change
A2.2   Affect: Cause/Connected
A3   Being
A4   Classification
A4.1   Generally kinds, groups, examples
A4.2   Particular/general; detail
A5   Evaluation
A5.1   Evaluation: Good/bad
A5.2   Evaluation: True/false
A5.3   Evaluation: Accuracy
A5.4   Evaluation: Authenticity
A6   Comparing
A6.1   Comparing: Similar/different
A6.2   Comparing: Usual/unusual
A6.3   Comparing: Variety
A7   Definite (+ modals)
A8   Seem
A9   Getting and giving; possession
A10   Open/closed; Hiding/Hidden; Finding; Showing
A11   Importance
A11.1   Importance: Important

A11.2   Importance: Noticeability
A12   Easy/difficult
A13   Degree
A13.1   Degree: Non-specific
A13.2   Degree: Maximizers
A13.3   Degree: Boosters
A13.4   Degree: Approximators
A13.5   Degree: Compromisers
A13.6   Degree: Diminishers
A13.7   Degree: Minimizers
A14   Exclusivizers/particularizers
A15 Safety/Danger
B THE BODY & THE INDIVIDUAL
B1   Anatomy and physiology
B2   Health and disease
B3   Medicines and medical treatment
B4   Cleaning and personal care
B5   Clothes and personal belongings
C ARTS & CRAFTS
C1   Arts and crafts
E EMOTIONAL ACTIONS, STATES & PROCESSES
E1   General
E2   Liking
E3   Calm/Violent/Angry
E4   Happy/sad
E4.1   Happy/sad: Happy
E4.2   Happy/sad: Contentment
E5   Fear/bravery/shock
E6   Worry, concern, confident
F FOOD & FARMING
F1   Food
F2   Drinks
F3   Cigarettes and drugs
F4   Farming & Horticulture

G GOVT. & THE PUBLIC DOMAIN
G1   Government, Politics & elections
G1.1   Government etc.
G1.2   Politics
G2   Crime, law and order
G2.1   Crime, law and order: Law & order
G2.2   General ethics
G3   Warfare, defence and the army; Weapons

H ARCHITECTURE, BUILDINGS, HOUSES & THE HOME
H1   Architecture, kinds of houses & buildings
H2   Parts of buildings
H3   Areas around or near houses
H4   Residence
H5   Furniture and household fittings

I MONEY & COMMERCE
I1   Money generally
I1.1   Money: Affluence
I1.2   Money: Debts
I1.3   Money: Price
I2   Business
I2.1   Business: Generally
I2.2   Business: Selling
I3   Work and employment
I3.1   Work and employment: Generally
I3.2   Work and employment: Professionalism
I4   Industry

K ENTERTAINMENT, SPORTS & GAMES
K1   Entertainment generally
K2   Music and related activities
K3   Recorded sound etc.
K4   Drama, the theatre & show business
K5   Sports and games generally
K5.1   Sports
K5.2   Games
K6   Children's games and toys

L LIFE & LIVING THINGS
L1   Life and living things
L2   Living creatures generally
L3   Plants

M MOVEMENT, LOCATION, TRAVEL & TRANSPORT
M1   Moving, coming and going
M2   Putting, taking, pulling, pushing, transporting &c.
M3   Movement/transportation: land
M4   Movement/transportation: water
M5   Movement/transportation: air
M6   Location and direction
M7   Places
M8   Remaining/stationary

N NUMBERS & MEASUREMENT
N1   Numbers
N2   Mathematics
N3   Measurement
N3.1   Measurement: General
N3.2   Measurement: Size
N3.3   Measurement: Distance
N3.4   Measurement: Volume
N3.5   Measurement: Weight
N3.6   Measurement: Area
N3.7   Measurement: Length & height
N3.8   Measurement: Speed
N4   Linear order
N5   Quantities
N5.1   Entirety; maximum
N5.2   Exceeding; waste
N6   Frequency etc.

O SUBSTANCES, MATERIALS, OBJECTS & EQUIPMENT
O1   Substances and materials generally
O1.1   Substances and materials generally: Solid
O1.2   Substances and materials generally: Liquid
O1.3   Substances and materials generally: Gas
O2   Objects generally
O3   Electricity and electrical equipment
O4   Physical attributes
O4.1   General appearance and physical properties
O4.2   Judgement of appearance (pretty etc.)
O4.3   Colour and colour patterns
O4.4   Shape
O4.5   Texture

O4.6 Temperature
P EDUCATION
P1 Education in general
Q LINGUISTIC ACTIONS, STATES & PROCESSES
Q1 Communication
Q1.1 Communication in general
Q1.2 Paper documents and writing
Q1.3 Telecommunications
Q2 Speech acts
Q2.1 Speech etc: Communicative
Q2.2 Speech acts
Q3 Language, speech and grammar
Q4 The Media
Q4.1 The Media: Books
Q4.2 The Media: Newspapers etc.
Q4.3 The Media: TV, Radio & Cinema
S SOCIAL ACTIONS, STATES & PROCESSES
S1 Social actions, states & processes
S1.1 Social actions, states & processes
S1.1.1 General
S1.1.2 Reciprocity
S1.1.3 Participation
S1.1.4 Deserve etc.
S1.2 Personality traits
S1.2.1 Approachability and Friendliness
S1.2.2 Avarice
S1.2.3 Egoism
S1.2.4 Politeness
S1.2.5 Toughness; strong/weak
S1.2.6 Sensible
S2 People
S2.1 People: Female
S2.2 People: Male
S3 Relationship
S3.1 Relationship: General
S3.2 Relationship: Intimate/sexual
S4 Kin
S5 Groups and affiliation
S6 Obligation and necessity
S7 Power relationship
S7.1 Power, organizing
S7.2 Respect
S7.3 Competition

S7.4 Permission
S8 Helping/hindering
S9 Religion and the supernatural
T TIME
T1 Time
T1.1 Time: General
T1.1.1 Time: General: Past
T1.1.2 Time: General: Present; simultaneous
T1.1.3 Time: General: Future
T1.2 Time: Momentary
T1.3 Time: Period
T2 Time: Beginning and ending
T3 Time: Old, new and young; age
T4 Time: Early/late
W THE WORLD & OUR ENVIRONMENT
W1 The universe
W2 Light
W3 Geographical terms
W4 Weather
W5 Green issues
X PSYCHOLOGICAL ACTIONS, STATES & PROCESSES
X1 General
X2 Mental actions and processes
X2.1 Thought, belief
X2.2 Knowledge
X2.3 Learn
X2.4 Investigate, examine, test, search
X2.5 Understand
X2.6 Expect
X3 Sensory
X3.1 Sensory: Taste
X3.2 Sensory: Sound
X3.3 Sensory: Touch
X3.4 Sensory: Sight
X3.5 Sensory: Smell
X4 Mental object
X4.1 Mental object: Conceptual object
X4.2 Mental object: Means, method
X5 Attention
X5.1 Attention
X5.2 Interest, boredom, excited, energetic
X6 Deciding
X7 Wanting; planning; choosing

X8	Trying	Z0	Unmatched proper noun
X9	Ability	Z1	Personal names
X9.1	Ability: Ability, intelligence	Z2	Geographical names
X9.2	Ability: Success and failure	Z3	Other proper names
Y SCIENCE & TECHNOLOGY		Z4	Discourse Bin
Y1	Science and technology in general	Z5	Grammatical bin
		Z6	Negative
Y2	Information technology and computing	Z7	If
		Z8	Pronouns etc.
Z NAMES & GRAMMATICAL WORDS		Z9	Trash can
		Z99	Unmatched

APPENDIX C

TYPE LIST OF THE WORDS IN THE COURSEBOOK ACCORDING TO K1

WORDS, K2 WORDS, OFF-LIST WORDS AND AWL WORDS

K1 TYPES (1-1000)

1k types: [families 709 : types 1273 : tokens 22286] a_[647] able_[1] about_[231] above_[13] accept_[12] accepting_[4] accepts_[4] according_[7] account_[4] across_[1] act_[4] action_[2] actions_[3] active_[3] activities_[37] activity_[68] actually_[2] add_[19] address_[5] advanced_[1] adventure_[9] adventures_[20] adventurous_[1] after_[41] again_[29] against_[4] age_[1] agency_[2] ago_[1] agree_[2] air_[6] all_[45] allow_[2] almost_[2] alone_[2] along_[1] already_[1] also_[28] always_[26] am_[94] among_[3] amount_[3] amounts_[1] an_[94] ancient_[3] and_[837] animals_[1] another_[4] answer_[58] answering_[3] answers_[16] any_[21] anyone_[1] anything_[16] anyway_[1] anywhere_[2] appears_[1] apply_[1] are_[309] arms_[4] around_[17] arrive_[2] art_[1] articles_[2] arts_[1] as_[68] ask_[55] asked_[3] asking_[4] asks_[1] at_[185] august_[5] away_[9] back_[16] bad_[10] ball_[2] bank_[1] base_[11] basic_[2] basis_[1] be_[61] bear_[1] beautiful_[4] beauty_[1] became_[2] because_[54] become_[5] becoming_[1] bed_[12] bedrooms_[2] beds_[2] before_[15] began_[3] begin_[2] beginning_[1] begun_[1] being_[4] believe_[3] belong_[1] below_[69] best_[16] better_[4] between_[13] big_[8] bigger_[1] biggest_[2] bill_[17] bird_[1] birds_[1] black_[2] blue_[9] board_[3] boat_[1] bodies_[1] body_[1] book_[6] books_[5] both_[4] bought_[2] box_[33] boxes_[29] bread_[12] break_[5] breaking_[2] bridge_[3] bridges_[1] bright_[4] bring_[10] bringing_[1] broke_[1] broken_[3] brother_[29] brothers_[7] brought_[3] build_[2] building_[7] buildings_[10] built_[5] burn_[1] burning_[2] but_[68] buy_[4] buying_[1] by_[31] call_[20] called_[6] calling_[14] calls_[3] came_[1] can_[255] cans_[1] capital_[7] car_[3] care_[20] careful_[1] carried_[1] castle_[1] catching_[8] cause_[2] caused_[6] causes_[6] causing_[1] center_[4] central_[2] centre_[1] certain_[1] change_[6] changed_[1] changes_[1] changing_[2] children_[6] choice_[1] choices_[4] choose_[7] chose_[1] chosen_[1] church_[3] churches_[1] circle_[5] circular_[2] cities_[3] city_[30] class_[34] clear_[5] close_[3] closer_[2] closest_[1] coast_[5] cold_[5] color_[2] come_[42] comes_[1] coming_[5] common_[8] complete_[29] completely_[1] concern_[2] concerns_[3] consider_[3] contain_[3] containing_[1] contains_[1] content_[1] contents_[1] continuously_[2] corn_[2] cost_[2] could_[21] count_[4] countries_[6] country_[2] course_[2] cover_[15] covered_[1] covering_[1] currently_[1] cut_[18] dad_[6] daily_[22] danger_[3] dangerous_[18] date_[3] day_[39] days_[1] deal_[1] deals_[1] dear_[4] death_[1] deaths_[1] december_[2] decide_[11] decided_[3] deep_[2] degrees_[1] depend_[1] depth_[1] describe_[14] describing_[5] description_[4] descriptions_[4] descriptive_[1] desert_[2] destroy_[5] destructive_[1] details_[3] develop_[3] developed_[6] did_[22] die_[2] died_[5] difference_[1] differences_[1] different_[22] difficult_[3] directions_[2] director_[1]

discover_[5]   discovered_[4]   discoveries_[3]   discovering_[1]   discovery_[5]
dislike_[2]   dislikes_[10]   do_[359]   doctor_[1]   does_[53]   dog_[4]   doing_[48]
dollar_[2]   don_[1]   done_[2]   down_[13]   draw_[5]   drawn_[1]   drew_[1]   drink_[6]
drinking_[3]   drinks_[3]   drive_[8]   driven_[1]   drives_[1]   driving_[1]   drop_[3]
drove_[1]   dry_[6]   duty_[1]   each_[17]   early_[12]   earth_[14]   earthquake_[27]
easier_[7]   easily_[1]   east_[3]   easy_[7]   eat_[12]   eating_[3]   effect_[2]   effectively_[2]
effects_[2]   efficiency_[1] egg_[4] eggs_[8] eight_[1] eightieth_[1] eleven_[1] else_[1]
empire_[2]   end_[5]   endless_[1]   english_[8]   enjoy_[24]   enjoying_[1]   enough_[8]
entering_[2]   entries_[2]   even_[5]   evening_[17]   evenings_[3]   event_[3]   events_[5]
ever_[7]   every_[26]   everyday_[2]   everyone_[1]   everything_[2]   everywhere_[5]
example_[71] examples_[11] except_[1] exchange_[17] exchanging_[1] exercise_[3]
experience_[5]   experienced_[1]   experiment_[2]   experiments_[4]   explain_[3]
explains_[1]   express_[17]   expressing_[3]   expression_[1]   expressions_[14]   eye_[9]
face_[17]   faces_[1]   fair_[5]   fall_[5]   fallen_[1]   falling_[1]   falls_[4]   familiar_[1]
families_[2]   family_[29]   famous_[3]   far_[3]   fast_[13]   faster_[1]   father_[20]
favorite_[32]   feed_[6]   feel_[7]   feeling_[4]   feelings_[8]   fell_[1]   felt_[2]   few_[6]
fifteen_[2]   filled_[20]   find_[22]   fine_[1]   finest_[1]   finish_[2]   finished_[3]   fire_[3]
fires_[1]   first_[50]   fish_[6]   fishes_[2]   fit_[2]   fits_[1]   five_[7]   fix_[2]   fixing_[3]
flew_[2]   floor_[11]   floors_[3]   flowers_[2]   fly_[4]   follow_[43]   following_[4]
food_[12] for_[223] forces_[19] forest_[6] forget_[4] form_[3] formed_[1] forms_[3]
found_[2]   four_[14]   fourteen_[9]   fourth_[1]   free_[11]   freedom_[2]   fresh_[2]
friday_[6]   friend_[24]   friendly_[4]   friends_[56]   friendship_[25]   friendships_[1]
from_[41]   front_[1]   full_[1]   furniture_[2]   further_[2]   future_[17]   game_[3]
games_[17]   garden_[4]   gases_[3]   gate_[2]   gather_[3]   gave_[2]   general_[1]
generally_[3]   gently_[1]   get_[39]   gets_[4]   getting_[2]   gift_[1]   girl_[2]   give_[24]
given_[6]   gives_[2]   giving_[3]   glad_[1]   glass_[10]   glasses_[3]   go_[97]   god_[2]
goes_[6]   going_[90]   gone_[3]   good_[34]   goodbye_[6]   got_[3]   gotten_[1]   great_[36]
greatest_[1]   green_[22]   grew_[1]   ground_[6]   group_[1]   groups_[8]   grow_[1]
grown_[1]   had_[7]   half_[8]   hand_[2]   hands_[5]   handwriting_[1]   hang_[22]
hanging_[3]   happen_[3]   happened_[2]   happens_[2]   happy_[4]   hard_[9]   harder_[2]
hardest_[1]   hardly_[6]   has_[31]   have_[130]   having_[12]   he_[87]   head_[4]   hear_[7]
heard_[3]   heart_[2]   heat_[10]   heating_[1]   heavy_[4]   held_[2]   help_[53]   helped_[1]
helps_[3]   her_[51]   here_[12]   high_[12]   highest_[3]   him_[21]   his_[43]   historic_[16]
historical_[3]   history_[3]   hold_[6]   holds_[1]   home_[41]   homeless_[2]   homes_[5]
hope_[12]   hot_[14]   hour_[8]   hours_[10]   house_[16]   how_[118]   however_[1]
human_[2]   i_[534]   idea_[9]   ideas_[14]   if_[51]   important_[16]   in_[541]   inch_[1]
include_[7]   including_[3]   inclusive_[2]   increase_[2]   increasing_[1]   instead_[1]
interest_[1]   interested_[8]   interesting_[12]   interests_[1]   into_[28]   is_[495]   it_[362]
its_[18]   itself_[1]   join_[12]   joined_[1]   joining_[2]   just_[16]   keep_[12]   keeps_[2]
kept_[2]   kind_[13]   kinds_[2]   knew_[2]   know_[16]   knowledge_[3]   known_[2]
knows_[1]   land_[2]   lands_[1]   language_[17]   languages_[4]   large_[11]   largest_[4]
last_[7]   lasts_[1]   late_[10]   later_[8]   lay_[1]   lead_[1]   leading_[1]   learn_[3]
learning_[1]   least_[9]   leave_[32]   leaves_[1]   leaving_[1]   led_[2]   left_[4]   lengths_[1]
less_[6]   let_[14]   letter_[16]   letters_[3]   lie_[2]   life_[26]   lifting_[1]   light_[14]
lightly_[1]   lights_[1]   like_[156]   likely_[1]   likes_[12]   limit_[1]   limited_[2]   line_[7]
lines_[5]   listen_[44]   listener_[1]   listening_[14]   little_[47]   live_[10]   lives_[5]
living_[1]   local_[5]   long_[6]   longer_[2]   look_[22]   looking_[3]   lose_[1]   lost_[2]

love_[24] lovely_[1] loves_[1] low_[4] machine_[1] made_[8] main_[11] make_[93] makes_[8] making_[8] man_[5] manufacturer_[1] manufacturers_[2] many_[48] mark_[1] market_[2] mass_[2] master_[1] material_[3] materials_[2] matter_[1] matters_[1] may_[9] maybe_[3] me_[40] mean_[29] meaning_[1] meaningful_[2] meanings_[4] means_[7] meant_[2] measured_[2] meet_[17] meeting_[1] meets_[1] member_[10] members_[3] memory_[3] men_[1] mentioned_[1] met_[2] metal_[4] middle_[6] midnight_[1] milk_[9] million_[13] millions_[1] mind_[4] mines_[1] minute_[4] minutes_[27] missed_[1] missing_[5] modern_[9] mom_[15] moment_[17] monday_[6] money_[1] month_[8] months_[2] moon_[1] more_[51] morning_[16] most_[39] mostly_[9] mother_[24] motor_[5] mountain_[5] mountains_[6] mouth_[1] move_[7] moved_[1] movement_[4] movements_[2] moves_[2] moving_[1] mr_[5] mrs_[4] much_[13] music_[26] musical_[1] must_[8] my_[219] myself_[4] name_[29] named_[1] names_[15] national_[6] nations_[1] natural_[32] nature_[8] near_[5] nearly_[6] necessary_[5] need_[11] needed_[7] neighborhood_[4] never_[19] new_[17] news_[13] newspaper_[5] newspapers_[1] next_[11] night_[6] nine_[2] no_[19] none_[1] north_[3] north-west_[1] not_[136] note_[17] notes_[4] nothing_[3] november_[1] now_[30] nowadays_[1] number_[43] numbers_[15] object_[1] objects_[2] observe_[1] of_[375] off_[28] offer_[2] offers_[5] official_[2] officials_[1] often_[35] oh_[7] oil_[10] old_[7] oldest_[1] on_[227] once_[13] one_[44] ones_[15] only_[4] onto_[2] open_[4] opening_[1] operate_[1] operating_[1] opinion_[3] opinions_[2] opportunities_[1] or_[122] order_[4] ordered_[2] organization_[1] organize_[1] organizing_[6] other_[18] others_[9] our_[22] ourselves_[1] out_[59] outdoors_[1] outside_[9] outstanding_[2] over_[21] own_[17] page_[1] pages_[1] paid_[2] paper_[11] part_[20] particular_[1] particularly_[1] parties_[2] parts_[10] party_[20] past_[5] pay_[3] people_[79] per_[5] person_[14] personal_[7] picture_[5] pictures_[4] piece_[10] place_[20] places_[17] plain_[2] plan_[8] planned_[1] planning_[2] plans_[12] plants_[6] play_[41] players_[1] playing_[8] plays_[2] please_[44] point_[13] points_[22] popular_[23] popularity_[1] population_[2] possible_[6] post_[3] poster_[13] posters_[4] powered_[6] powerful_[4] prepare_[9] prepared_[2] present_[3] presentation_[1] presented_[2] prices_[1] problem_[10] problems_[5] producing_[2] product_[2] production_[10] products_[2] progress_[10] promise_[1] proper_[1] protect_[4] protection_[1] provided_[1] public_[3] put_[48] putting_[4] quality_[3] quantities_[1] question_[7] questions_[76] quite_[2] race_[4] racing_[17] raise_[2] raised_[1] ran_[1] rather_[7] reach_[1] reached_[1] reaches_[1] read_[97] reading_[12] readings_[1] reads_[1] ready_[5] real_[9] really_[26] reason_[2] reasons_[22] receive_[3] received_[1] receiver_[3] recent_[1] recognize_[4] recognizes_[1] record_[2] recorded_[5] recording_[22] red_[3] reduce_[5] refuse_[13] refuses_[4] refusing_[3] related_[6] relatives_[2] religions_[1] remain_[2] remember_[2] reply_[2] report_[1] reports_[3] respect_[2] rest_[3] result_[4] results_[1] rich_[1] ride_[6] rider_[3] riders_[1] riding_[2] right_[6] rise_[1] river_[1] rivers_[2] riverside_[3] road_[6] rock_[13] rocky_[1] rode_[1] roll_[5] rolling_[3] room_[17] rooms_[1] rose_[1] round_[1] rounds_[2] rule_[2] rules_[10] run_[4] runner_[2] runners_[1] running_[14] safe_[4] safer_[2] safety_[3] said_[8] salt_[17] same_[6] sat_[2] saturday_[19] saturdays_[1] save_[8] saving_[5] saw_[1] say_[28] says_[1] scene_[2] scenery_[4] school_[72] schools_[2] science_[54] scientific_[9] scientist_[3] scientists_[3] sea_[9] seaside_[2] season_[1] seat_[4] second_[2]

seconds_[1] secret_[2] see_[34] seeing_[1] seem_[1] seems_[1] seen_[1] send_[19]
sender_[3] sending_[1] sent_[2] separate_[6] serve_[5] service_[4] services_[1]
serving_[3] set_[8] settled_[1] seven_[5] several_[5] shake_[4] shaking_[3] shall_[1]
shape_[4] share_[57] sharing_[1] she_[68] shook_[3] short_[47] shortage_[3]
shorter_[3] should_[16] show_[12] showed_[3] showing_[1] shown_[1] shows_[3]
side_[3] sides_[3] sign_[6] signs_[1] silver_[2] simple_[55] since_[1] sing_[2]
singer_[4] sister_[13] sisters_[2] sit_[2] sits_[1] situations_[1] six_[3] size_[9]
sky_[4] sleep_[2] small_[6] smaller_[1] smiling_[1] snow_[9] so_[42] social_[10]
soft_[1] softened_[1] some_[48] someone_[15] something_[22] sometimes_[31]
song_[2] songs_[3] soon_[9] sort_[1] sound_[1] sounded_[1] sounds_[3] south_[4]
southern_[1] space_[6] spaces_[1] speak_[13] speakers_[2] speaking_[14] special_[8]
speed_[5] speeds_[1] spend_[7] spends_[1] spent_[2] spoke_[1] spoken_[21]
spread_[4] springs_[1] square_[3] stage_[1] stand_[2] standard_[4] stands_[1] star_[1]
start_[13] started_[4] starts_[7] state_[9] statements_[20] stay_[12] stayed_[1]
steel_[2] step_[4] steps_[2] still_[2] stock_[1] stood_[2] stop_[5] stopping_[2]
stops_[3] store_[3] stores_[5] storing_[3] stranger_[5] strangers_[1] street_[3]
streets_[1] strike_[1] stroke_[1] strong_[6] stronger_[1] struck_[1] student_[37]
students_[5] studies_[1] study_[32] studying_[2] subject_[1] substance_[3]
successfully_[3] such_[20] suggestion_[1] suggestions_[5] summer_[5] summers_[2]
sunday_[11] sundays_[4] sunny_[1] support_[4] supporting_[2] suppose_[1]
supposed_[3] sure_[7] surely_[1] surface_[4] surprise_[1] sweet_[1] system_[1]
systems_[1] table_[9] take_[70] taken_[4] takes_[2] taking_[4] talk_[16] talking_[10]
talks_[2] teach_[1] teacher_[15] teachers_[4] tell_[16] tells_[2] ten_[3] term_[3]
test_[8] tested_[2] testing_[1] than_[41] that_[82] the_[1779] their_[36] them_[46]
themselves_[2] then_[50] there_[71] these_[31] they_[109] thing_[7] things_[26]
think_[49] third_[3] thirds_[2] this_[59] those_[1] thought_[2] thousand_[1]
thousands_[1] three_[22] through_[9] throughout_[1] thursday_[4] time_[42]
times_[11] to_[637] today_[12] together_[13] told_[2] too_[14] took_[6] top_[6]
total_[3] touch_[8] town_[2] trader_[1] traders_[1] train_[2] trainers_[1] travel_[3]
traveled_[1] travelling_[3] tree_[1] trees_[3] tried_[3] true_[22] truly_[4] try_[10]
trying_[3] tuesday_[4] turn_[24] turns_[4] twentieth_[2] twice_[8] two_[31] type_[2]
types_[6] uncontrolled_[1] under_[19] understand_[45] understanding_[1]
understood_[2] unit_[194] united_[2] unknown_[2] unless_[1] unlimited_[3] until_[9]
unusual_[3] up_[51] upside_[2] us_[26] use_[108] used_[1] users_[8] uses_[10]
using_[16] usually_[65] various_[5] very_[53] viewing_[1] visit_[22] visiting_[9]
voice_[8] wait_[6] waiting_[1] walk_[16] walking_[5] walks_[2] wall_[3] walls_[3]
want_[9] wants_[1] was_[38] watch_[24] watching_[9] water_[57] watering_[1]
waters_[3] wave_[5] waves_[3] way_[3] ways_[2] we_[96] wear_[3] wednesday_[7]
week_[16] weekend_[10] weekends_[6] weeks_[1] welcome_[2] welcoming_[1]
well_[48] wells_[1] went_[5] were_[11] west_[4] what_[237] whatever_[1]
when_[70] whenever_[2] where_[36] which_[19] while_[16] white_[10] who_[43]
whole_[1] why_[46] wide_[3] wife_[1] will_[43] willingness_[2] wind_[4]
window_[1] windows_[5] winds_[1] winter_[2] winters_[3] wish_[1] wishes_[2]
with_[255] within_[2] without_[2] woman_[1] women_[1] wonder_[2] wonderful_[1]
wonders_[2] wood_[1] word_[10] words_[38] wore_[1] work_[60] worked_[1]
worker_[1] workers_[2] working_[9] world_[43] worn_[1] worth_[1] would_[83]
write_[138] writer_[1] writes_[1] writing_[13] written_[2] wrong_[1] wrote_[1]

year_[11] years_[9] yes_[28] yesterday_[2] you_[603] young_[4] your_[364] yourself_[11]

K2 TYPES (1001-2000)

2k types: [370:530:2677] advertisement_[1] advertisements_[2] afraid_[2] afternoon_[25] afternoons_[2] aim_[1] apologies_[1] apologize_[2] apology_[1] argue_[1] arrangements_[1] asleep_[1] astonishing_[2] attend_[2] attending_[2] attention_[2] attraction_[3] attractions_[16] attracts_[1] aunt_[3] avoid_[1] bag_[1] bags_[2] bake_[9] baking_[6] band_[9] barber_[1] bath_[5] baths_[3] beat_[1] beaten_[1] behavior_[1] bell_[5] berries_[2] berry_[2] bicycle_[5] billion_[5] birthday_[5] block_[1] blocks_[3] boil_[9] boiled_[1] boils_[1] border_[3] bottles_[1] bottom_[1] bowl_[11] bowling_[2] breakfast_[18] brown_[4] brush_[8] brushing_[7] buried_[2] bus_[8] busy_[7] butter_[12] button_[1] cake_[2] cakes_[2] camera_[4] camp_[1] camping_[1] card_[1] cards_[2] cat_[1] cattle_[1] cave_[2] caving_[6] century_[2] ceremony_[1] chain_[4] chairs_[1] cheap_[3] check_[68] checking_[10] checks_[1] cheers_[2] cheese_[9] chicken_[4] clean_[18] cleaning_[4] climbers_[2] climbing_[7] clock_[11] cloth_[2] clothes_[5] club_[3] coffee_[4] collection_[1] combination_[3] combinations_[1] comfortable_[1] compare_[10] comparing_[3] comparison_[1] comparisons_[7] competition_[2] competitions_[2] connect_[9] connected_[3] connecting_[1] connection_[20] conscious_[2] convenient_[1] conversation_[20] conversations_[7] cook_[16] cooker_[2] cooking_[24] cooks_[4] cool_[6] copy_[4] correct_[28] cousin_[4] cousins_[1] crack_[1] crash_[1] crashing_[1] creatures_[1] crops_[2] crushed_[1] cup_[4] cupboard_[1] cups_[1] cure_[1] damage_[5] damaged_[1] damp_[2] delicate_[2] desk_[6] dictionary_[13] dinner_[12] dirty_[3] discuss_[2] discussion_[3] discussions_[2] diseases_[2] dish_[13] dishes_[29] diving_[11] double_[4] drums_[1] during_[13] dust_[8] dusting_[3] electric_[8] electricity_[4] empty_[5] emptying_[1] engine_[4] engineer_[2] engineering_[2] entertaining_[10] entire_[1] especially_[2] excited_[3] exciting_[11] excuse_[8] excuses_[11] explore_[4] exploring_[2] explosion_[1] extra_[1] extraordinary_[2] extreme_[18] extremely_[3] false_[15] fan_[1] fans_[2] fashion_[3] fasten_[1] feet_[2] film_[1] fingers_[1] firmly_[1] flag_[1] flash_[2] flat_[1] flood_[4] flooding_[3] floods_[2] flour_[11] fold_[5] formal_[2] frame_[4] frequency_[4] fried_[5] frighten_[1] frightened_[1] frozen_[1] fruit_[4] fry_[9] frying_[4] fun_[23] funny_[1] gaps_[1] grammar_[2] grand_[4] grandmother_[1] grandparents_[6] greased_[1] greet_[4] guess_[7] guesses_[1] guests_[1] habit_[2] habits_[13] hair_[1] harbor_[1] harmful_[1] hat_[1] hate_[3] health_[5] healthier_[1] healthy_[1] hello_[41] hey_[4] hi_[28] hidden_[1] hit_[7] hits_[1] holidays_[1] honest_[1] hospitals_[1] hotel_[2] hotels_[1] hungry_[1] hunt_[1] hurt_[1] ice_[2] imagine_[10] immediately_[3] improve_[2] improved_[2] improving_[2] inform_[5] informal_[1] information_[67] inside_[2] instruments_[3] intended_[1] invent_[4] invented_[9] invention_[9] inventions_[7] inventor_[5] inventors_[4] invitation_[27] invitations_[9] invite_[6] invited_[1] invites_[3] inviting_[2] irregular_[1] island_[7] islands_[2] journey_[1] juice_[3] jump_[4] jumping_[11] jumps_[1] key_[1]

kilometres_[1] kitchen_[3] knives_[1] lean_[1] leans_[1] legs_[1] lesson_[60] lessons_[9] liberty_[2] list_[13] lists_[2] load_[2] loads_[1] loaf_[4] lock_[1] log_[8] lonely_[1] lot_[6] lots_[7] loud_[4] lunch_[10] mail_[2] management_[1] map_[1] maps_[1] match_[42] meal_[4] meals_[3] meat_[5] mechanical_[5] medicine_[6] melt_[3] melted_[1] melts_[4] message_[23] messages_[23] mild_[2] milligram_[1] mistakes_[1] mix_[5] mixed_[2] mixing_[4] mixture_[7] model_[1] motion_[4] mouse_[1] mysterious_[2] neat_[1] neck_[1] net_[18] nice_[4] noise_[2] noon_[2] noun_[18] nuts_[2] ocean_[2] oceans_[1] orange_[3] organ_[2] originally_[1] outline_[5] package_[1] packed_[1] packs_[1] pain_[1] pair_[1] pairs_[18] pale_[2] pan_[16] pardon_[2] parents_[24] park_[4] parks_[1] passages_[6] patient_[1] pen_[3] perfect_[2] perfected_[1] perform_[2] performing_[3] pet_[1] phone_[20] phones_[5] photographs_[2] pick_[6] pin_[3] pipe_[1] plate_[1] pocket_[19] poem_[10] poems_[1] polish_[1] pot_[1] potted_[2] pour_[7] powder_[5] practical_[4] practice_[17] practicing_[1] precious_[3] prefer_[41] preference_[5] preferences_[11] pretend_[4] print_[2] printed_[1] prize_[7] prizes_[2] probable_[1] program_[3] programs_[1] pronounce_[1] pronounces_[1] push_[1] puzzle_[4] quickly_[5] quiet_[1] rain_[5] rarely_[3] ray_[3] regular_[7] regularly_[4] repair_[1] repeat_[8] request_[1] requests_[1] rescue_[2] rescued_[1] responsibilities_[17] responsibility_[3] responsible_[20] restaurant_[1] rice_[11] risk_[2] roast_[6] rob_[5] rubbish_[1] rude_[1] rush_[1] sad_[4] samples_[1] sandy_[2] sauce_[5] scale_[4] screen_[10] screw_[3] search_[15] searching_[1] seeded_[1] self_[10] sentence_[1] sentences_[63] sheep_[1] sheet_[1] shelves_[6] shirt_[2] shirts_[2] shocking_[1] shoes_[5] shopping_[20] shops_[3] shower_[8] skills_[1] slave_[1] slide_[2] slope_[1] slower_[1] socks_[2] soil_[1] solve_[2] sorry_[15] soup_[1] spelling_[2] spin_[1] spinning_[1] sport_[10] sports_[57] stairs_[1] steam_[4] steamed_[1] steep_[1] stick_[2] stir_[8] stirring_[2] storm_[2] stove_[3] stretches_[1] sudden_[2] suddenly_[1] sugar_[8] suitable_[3] swear_[1] sweep_[2] sweeping_[1] swimming_[2] sympathy_[3] tall_[6] tallest_[3] tap_[4] taps_[3] tea_[6] teeth_[13] telephone_[6] temperature_[2] temperatures_[1] tender_[1] thanks_[20] theatres_[2] thickly_[2] ticket_[1] tickets_[5] tidy_[5] tidying_[2] tire_[4] tired_[2] tiring_[2] title_[3] titled_[1] titles_[1] tomorrow_[17] tonight_[9] tooth_[1] tourist_[26] tourists_[4] tours_[1] towel_[1] tower_[12] towers_[2] toys_[1] tracking_[2] translate_[1] translates_[1] trapped_[3] treated_[1] treatment_[1] trip_[3] tube_[4] tubes_[1] typical_[2] uncle_[4] uncomfortable_[2] universe_[5] untidy_[1] upper_[1] verb_[3] verbs_[2] violent_[4] warm_[5] warmed_[1] warming_[4] wash_[14] washing_[6] waste_[6] wasting_[3] weather_[10] weed_[2] weigh_[1] weighs_[1] wet_[1] wheel_[4] wheels_[4] wire_[1] wires_[1] witness_[1] worming_[1] worried_[1] worry_[1] worst_[1]

ACADEMIC WORD TYPES

AWL types: [141:211:895] access_[5] accesses_[1] accommodation_[4] achievement_[2] achievements_[6] adjust_[3] adults_[1] affect_[2] affected_[1] affects_[1] aid_[4] aids_[2] alternative_[1] appropriate_[17] approximately_[8]

area_[4] areas_[1] assess_[10] assessment_[12] assistant_[1] attachment_[4] attitudes_[3] authorities_[3] authority_[1] available_[8] awareness_[3] brief_[1] capable_[3] category_[1] challenge_[1] challenging_[10] chart_[18] charts_[1] chemical_[4] clarification_[10] clarify_[3] classical_[3] co-operate_[1] code_[1] collapse_[5] comment_[2] communicate_[3] communicating_[2] communication_[32] compensation_[7] computer_[39] computers_[4] computing_[3] concentration_[1] concept_[3] conduct_[1] conducting_[1] confirm_[2] consists_[2] construct_[1] constructed_[1] construction_[3] contact_[3] context_[1] create_[4] cultural_[2] culture_[2] cultures_[2] currency_[2] cycle_[11] cycling_[17] definitely_[2] definition_[3] definitions_[8] depressed_[1] design_[12] designed_[8] designer_[3] device_[2] devices_[2] display_[9] displays_[1] disposable_[1] distributions_[1] diversity_[1] dominant_[1] dominates_[1] drama_[2] dramas_[1] economy_[1] energetic_[3] energy_[5] environment_[1] environmental_[3] equipment_[1] exhibit_[1] exhibition_[3] exhibitions_[1] expert_[3] experts_[1] features_[1] fee_[2] file_[1] finalized_[2] finally_[12] focus_[1] function_[1] global_[5] grade_[1] highlights_[2] identify_[9] identities_[2] identity_[1] impact_[2] injured_[5] injuries_[3] injury_[1] innovation_[2] innovations_[2] interaction_[12] interactive_[2] investigator_[2] involve_[4] involving_[1] item_[10] items_[2] journal_[1] labelled_[1] linked_[1] located_[3] location_[5] maintain_[1] major_[3] medical_[3] medium_[2] method_[1] methods_[1] mode_[1] monitor_[2] motivation_[1] negative_[1] networking_[8] networks_[2] normal_[1] occur_[2] occurring_[1] occurs_[1] paragraph_[34] paragraphs_[4] partner_[39] period_[4] physical_[1] physically_[3] predictable_[1] predictions_[5] principle_[1] principles_[1] priority_[2] process_[14] processed_[1] project_[16] projects_[1] range_[1] reactions_[1] region_[3] regional_[2] regions_[2] register_[4] registration_[2] remove_[2] research_[3] researchers_[1] residents_[1] respond_[6] responding_[1] response_[3] responses_[2] revolutionary_[1] role_[4] roles_[1] seek_[1] sequencing_[2] similar_[2] similarities_[1] site_[5] sites_[19] somewhat_[1] source_[1] sources_[1] specific_[5] specifications_[1] stable_[3] strategies_[7] structure_[8] structures_[5] style_[3] summary_[1] symbol_[2] symbols_[1] task_[2] tasks_[2] team_[2] technical_[1] technology_[8] text_[64] texts_[18] theories_[1] thesis_[3] topic_[4] topics_[4] traditional_[6] transportation_[3] unique_[3] variations_[2] vehicle_[1] volunteer_[2]

OFF-LIST WORD TYPES

OFF types: [?:1032:3557] a&#number;a_[1] abbasid_[1] abbreviation_[1] abbreviations_[4] abby_[4] absorb_[1] absorbs_[2] accelerate_[1] accelerating_[1] accessory_[2] activate_[1] activites_[1] ad_[1] ada_[1] adams_[2] addict_[4] addicts_[8] adeline_[1] adjectives_[4] adrenalin_[7] adress_[2] adverbs_[1] afghan_[3] afghanistan_[3] africa_[2] aftershock_[1] aftershocks_[1] agra_[1] ahmet_[5] aircraft_[8] alarms_[1] albert_[2] alexander_[4] algorithm_[1] alice_[1] allbright_[1] alley_[1] ally_[5] almonds_[2] alphabet_[1] aluminum_[1] amanda_[3] amazing_[2] america_[3] american_[2] amin_[1] ana_[1] anatolia_[2] anatolian_[1]

angela_[1] animated_[2] animation_[2] ankle_[1] ankles_[1] antalya_[3] antenna_[1] antony_[1] anzac_[1] apartment_[1] apologizing_[4] appetite_[1] appliances_[1] arab_[1] archimedes_[4] architect_[2] architectural_[3] architecture_[5] archive_[1] argentina_[6] ashley_[6] asia_[1] assyrians_[1] astronomy_[2] audio_[1] australia_[8] australians_[3] avalanche_[5] ave_[1] award_[2] awards_[1] ay&#number_[4] babbage_[5] babylonians_[1] backyards_[1] bacteria_[2] badminton_[1] balloon_[2] banana_[2] barbara_[1] barbecue_[3] barbeque_[2] barbeques_[1] basics_[1] basildon_[1] basketball_[2] bathroom_[7] bbl_[1] beach_[3] beaches_[3] bedesten_[1] beef_[6] belinda_[2] ben_[4] benjamin_[3] beshbarmak_[3] betsy_[4] betty_[2] bike_[35] biker_[1] bikes_[6] biking_[4] bin_[4] bins_[1] birmingham_[1] biscuit_[1] blend_[1] blizzard_[1] blog_[1] bnumbern_[2] bob_[26] bon_[1] bonjour_[1] bookcases_[1] booklet_[5] boring_[6] bosphorus_[1] bossaball_[3] boulder_[1] brackets_[3] brake_[8] brakes_[1] breakthrough_[2] brian_[4] britain_[1] british_[2] brochure_[19] brochures_[5] browser_[4] bruce_[3] buddies_[2] buddy_[1] budget_[2] bug_[1] bulb_[11] bulbs_[4] bungee_[5] bunsen_[2] burj_[2] burner_[2] bursa_[3] burt_[1] bye_[19] byzantines_[1] cabbage_[6] cabbages_[1] cable_[17] caf_[4] caliphate_[1] caller_[3] calories_[3] campaign_[1] campaigns_[1] canada_[3] canadian_[1] cancer_[4] candidate_[1] candied_[1] canoeing_[7] cansu_[5] canteen_[1] capitalization_[2] careers_[1] carell_[1] caribbean_[6] carmen_[1] carpets_[2] carrot_[6] carrots_[4] carty_[1] casual_[2] cauliflower_[2] cd_[5] cds_[1] cell_[2] cells_[4] centimeters_[2] champion_[1] charles_[5] chat_[6] chats_[2] chatting_[1] checklist_[24] chemist_[5] chemistry_[1] chess_[1] chestnut_[1] chili_[2] chilly_[2] china_[5] chinese_[1] chip_[1] chips_[2] chit_[2] chocolate_[2] choir_[1] chop_[3] chopped_[3] chore_[5] chores_[39] chromosomes_[1] cilicia_[1] cilician_[1] civilian_[1] civilization_[1] civilizations_[3] clare_[3] classmate_[2] classmates_[36] classroom_[6] cleopartra_[1] cleopatra_[2] climate_[8] cloves_[1] clues_[2] cm_[3] cnm_[1] coach_[1] coconut_[3] colcannon_[3] colorado_[2] com_[3] comedies_[5] comedy_[4] communicator_[2] compact_[3] compass_[3] concert_[2] concerts_[1] conditioner_[2] cone_[2] connectors_[3] conserve_[1] continental_[1] convey_[2] cookies_[1] corns_[2] cosmonaut_[2] cosmos_[2] couch_[1] couches_[1] counselling_[2] countryside_[1] coupon_[1] crackers_[1] craters_[1] crawl_[3] crossroad_[1] crossword_[3] cruise_[1] cruz_[1] crystal_[1] cuisine_[5] culnumberr_[1] curie_[1] curtis_[14] custard_[2] dam_[1] dams_[1] dan_[30] daniel_[1] daniela_[4] danny_[8] darren_[1] dave_[6] debris_[4] delete_[5] delicious_[5] demirci_[1] denis_[5] dentist_[1] desktop_[4] dessert_[2] desserts_[4] destination_[7] destinations_[2] detergents_[1] diagnose_[2] diagnoses_[1] dial_[1] dialog_[2] dialogs_[1] diameter_[1] diane_[4] diaries_[1] diary_[4] dice_[3] diced_[1] digital_[1] disaster_[6] disasters_[15] disconnect_[2] discoverer_[3] discoverers_[2] discs_[1] dishwasher_[7] disk_[2] disorganized_[1] dna_[11] donna_[11] dossier_[14] dough_[8] download_[5] drought_[5] drug_[1] drugs_[1] duabi_[1] dubai_[1] dvd_[1] eco_[5] ecosystem_[1] ecosystems_[2] edison_[5] eggplant_[1] eggplants_[3] egyptian_[4] eiffel_[3] electronic_[1] elementary_[1] elevator_[4] ellen_[6] elmo_[2] email_[20] emails_[10] emergency_[1] emir_[1] emirates_[1] emma_[1] emperor_[3] endeavor_[1] eric_[4] eruption_[2] eruptions_[1] eshab_[1] eski&#number;ehir_[1] essex_[1] etc_[6] europe_[1] european_[2] everest_[2] exam_[3] excavation_[1] fantastic_[4] fascinating_[10] fatih_[1] fax_[2] ferris_[1]

festival_[1] fi_[9] flashfloods_[2] florets_[1] floured_[2] fluorescent_[5] flute_[2] focal_[1] footwear_[1] forever_[4] fossils_[1] france_[2] frank_[28] frankie_[2] franklin_[3] fred_[2] french_[5] gadget_[1] gale_[13] galileo_[4] garbage_[8] garlic_[1] gb_[1] gear_[5] gears_[1] genetic_[2] geneticist_[1] geneticists_[1] geographical_[1] geography_[1] geology_[1] george_[1] german_[2] giant_[2] gina_[5] gino_[2] glacial_[3] glaciares_[1] glacier_[3] glaciers_[3] glen_[1] gliding_[8] gonna_[1] graduation_[1] gravity_[1] greenhouse_[3] grill_[6] grocery_[7] guitar_[3] guten_[1] guys_[1] gym_[2] hadzicki_[2] hagia_[3] hairdresser_[1] handlebar_[4] harfield_[1] harire_[2] harvard_[1] hasan_[1] headphones_[3] hearth_[1] heights_[1] helen_[27] helmet_[4] henry_[4] hereexcept_[1] heritage_[1] hiking_[5] hillary_[1] himeji_[1] hip_[2] hittites_[1] hiv_[1] hobby_[1] homework_[18] honey_[4] hop_[2] hostel_[1] hostels_[1] household_[3] housework_[7] huge_[3] hugs_[1] humanity_[2] hurricane_[4] i&#number;kbebet_[1] icon_[1] impressive_[4] incredible_[1] india_[2] indian_[1] indonesia_[1] indonesian_[1] indoor_[1] infancy_[1] ingredients_[17] inhale_[1] inhaling_[1] inland_[1] intercultural_[3] interior_[1] internet_[112] interview_[6] irish_[2] ironing_[2] italian_[4] italy_[5] jack_[9] jahan_[1] jam_[11] jamaica_[9] jamaican_[1] james_[2] jane_[6] japan_[3] japanese_[4] jason_[12] jasonseville_[1] jasonvill_[2] jazz_[4] jeff_[1] jen_[1] jeremy_[6] jill_[2] jim_[1] joe_[7] jogging_[1] john_[1] jug_[1] jumpers_[3] junction_[1] junk_[2] justinian_[1] kabul_[2] kabuli_[2] karting_[1] kate_[19] katya_[2] kayaking_[10] kebabs_[1] kehf_[1] kemalpa&#number;a_[1] ken_[2] kensington_[1] kevin_[6] kevlar_[3] keyboard_[4] khan_[1] khufu_[1] kilometers_[6] kilos_[1] kingston_[1] kit_[3] km_[3] knead_[3] kremlin_[1] kwolek_[3] lab_[1] lain_[1] lamb_[4] lamington_[1] landline_[2] landlines_[1] landslide_[8] landslides_[1] lanka_[1] laptop_[1] laptops_[1] lasagna_[2] launch_[1] laundry_[4] leeks_[1] lentils_[2] leo_[3] lever_[4] lifestyle_[2] lightning_[1] linda_[8] liter_[1] liz_[1] lol_[1] london_[6] los_[1] lovelace_[1] madrassas_[1] magazine_[2] magazines_[2] magnitude_[3] mahal_[3] majestic_[2] maldives_[1] mall_[2] malls_[2] marc_[9] mardin_[8] marie_[1] marmara_[1] martial_[1] martin_[14] mary_[3] mashed_[2] massive_[4] math_[2] matt_[2] maya_[1] mcclintock_[1] meatball_[1] medial_[1] mediterranean_[5] megan_[1] mehmet_[4] memorable_[2] merhaba_[1] meryem_[1] mesopotamia_[1] messaging_[5] meteor_[2] meters_[5] mexico_[4] micro_[2] microchip_[2] microphone_[1] microscopes_[1] microscopic_[1] microwave_[3] mike_[19] mini_[3] minigolf_[1] miracle_[1] missouri_[1] ml_[1] mobile_[11] mock_[3] modem_[20] monastery_[3] mop_[4] mopping_[2] moreno_[1] mosaics_[1] moscow_[1] mosque_[11] mosques_[1] motorbike_[6] motorbikes_[3] motorcyclists_[1] movie_[26] movies_[34] mu&#number;la_[1] mudguard_[4] mudguards_[2] muffin_[2] muffins_[1] multicultural_[2] mumtaz_[1] museum_[11] museums_[1] mushroom_[1] mushrooms_[2] musicals_[1] nami_[1] nancy_[9] nanomedicine_[5] nanomolecules_[2] nanotechnology_[3] narrative_[1] narratives_[3] nelly_[1] nev&#number;ehir_[1] niagara_[3] nicole_[1] nina_[4] nobel_[5] noodles_[2] northwestern_[1] notebook_[2] number;kbebet_[2] number;neg_[1] number;rkka&#number;&#number_[1] number;skender_[1] number;stanbul_[5] numberd_[2] numbernite_[1] oakland_[2] oats_[1] obligation_[1] obligations_[7] offline_[4] ok_[37] okay_[8] oklahoma_[2] ol_[1] oliver_[1] olympic_[1] olympics_[1] onda_[11] onion_[9] onions_[2] online_[35] orchestra_[1]

organism_[1] organisms_[1] orzo_[4] ottoman_[2] outrageously_[1] oven_[6] overflow_[2] pacific_[2] packet_[1] pal_[2] palace_[3] palaw_[2] palm_[1] palms_[2] pancake_[1] pancakes_[3] panic_[2] pants_[2] par_[1] parachute_[9] paris_[3] parker_[2] parkour_[15] participle_[2] password_[4] pasta_[6] patents_[2] patients_[1] paul_[2] pavlova_[1] pc_[9] pcs_[1] pedal_[4] pedals_[1] peel_[3] peeled_[1] peer_[2] pepper_[8] peppers_[1] performer_[3] perito_[1] persian_[1] pete_[6] peter_[12] peterson_[2] pharaoh_[1] phonograph_[5] photo_[6] photos_[52] phrase_[1] phrases_[16] physicist_[2] physics_[1] picnic_[4] pie_[1] pilaf_[4] pisa_[2] pistachios_[2] pixy_[1] pizza_[10] planet_[4] plastic_[6] pollution_[3] polonium_[1] poly_[7] pop_[9] popping_[1] portfolio_[1] postcard_[4] postcards_[1] potato_[1] potatoes_[3] pre_[1] preheat_[1] printer_[1] prix_[1] processor_[2] professor_[1] programmer_[2] prophet_[1] prototype_[3] prototypes_[2] proverb_[1] province_[6] pudding_[1] punctuation_[2] pyramid_[4] pyramids_[1] quake_[2] questionnaire_[2] quiz_[2] qwerty_[1] rachael_[8] rackets_[1] radar_[1] radioactivity_[1] radium_[1] raffle_[1] raft_[2] rafting_[20] rainfall_[1] rap_[1] reacher_[1] rear_[1] recipe_[10] recipes_[3] recordings_[6] recreation_[1] recreational_[3] rectangle_[1] recycle_[5] recycling_[1] reinvent_[1] reset_[1] resort_[4] reuse_[4] rewrite_[1] richter_[4] rick_[1] ridiculous_[1] rimms_[1] rinse_[3] risen_[1] robert_[4] robots_[1] roly_[7] roman_[4] romance_[3] romances_[4] romans_[1] rosa_[2] rossi_[1] routine_[4] routines_[9] rudd_[1] russia_[1] russian_[1] safari_[1] salad_[1] salim_[1] sally_[12] sam_[20] samuel_[1] sanders_[4] sandra_[4] santa_[1] saucepan_[3] scan_[12] scared_[1] schmucks_[1] sci_[9] score_[4] scores_[3] scott_[1] scripts_[1] scuba_[2] seafood_[2] searles_[1] seeker_[4] seekers_[3] seljuk_[1] sembusek_[2] serials_[1] settings_[1] seville_[1] sezen_[1] shah_[1] shortages_[3] sibling_[1] siblings_[1] sided_[3] sightseeing_[3] skateboarding_[9] skating_[1] ski_[2] skillful_[2] skis_[1] skyscraper_[1] slice_[5] sliced_[6] slices_[4] smartphone_[2] smartphones_[1] sms_[1] snack_[5] snacks_[3] snowboard_[1] snowshoeing_[3] soware_[1] soccer_[14] socialize_[1] soda_[3] software_[3] sonya_[1] sophia_[3] southeastern_[1] souvenirs_[1] soviet_[1] spam_[5] span_[1] spanish_[4] sparks_[1] sparksadeline_[1] specimens_[1] spectacular_[3] spencer_[1] sponsor_[1] sponsors_[1] sprinkle_[1] sprinkler_[2] sri_[1] ssnumberne_[1] st_[1] stacy_[1] stadium_[6] statue_[2] steak_[2] stephanie_[3] steve_[4] sticky_[1] storage_[2] strawberry_[1] stunning_[1] sturdy_[4] suburb_[1] suburbs_[2] sultan_[1] sumatra_[1] sumela_[2] surf_[2] surfing_[1] surgeries_[2] surgery_[1] surname_[1] suzy_[5] swift_[4] switch_[1] synthetic_[2] sytem_[1] tablespoon_[5] tablespoons_[7] tablet_[6] tablets_[1] tag_[1] taiwan_[2] taj_[2] tanzania_[1] tarsus_[11] tasty_[1] taylor_[4] teaspoon_[4] technician_[2] teen_[22] teenager_[3] teenagers_[9] teens_[18] telecommunication_[1] telescope_[5] tennis_[10] tereshkova_[1] terrific_[4] texting_[1] thailand_[1] thames_[1] theater_[3] theaters_[1] thomas_[5] thriller_[2] thrillers_[7] tigers_[1] tim_[4] timetable_[2] tina_[8] tiny_[1] tissues_[1] toast_[8] toaster_[2] tom_[1] tomato_[8] tomb_[3] topkap&#number_[1] tornado_[7] touchscreen_[1] tourism_[25] tournament_[4] trabzon_[1] traffic_[1] translator_[1] trekking_[1] trendy_[4] tropical_[3] tsu_[1] tsunami_[8] tsunamis_[1] tubing_[4] turbine_[1] turkey_[16] turkish_[7] turkmen_[1] turkmenistan_[1] turks_[1] tutor_[1] tv_[16] tyler_[5] tyres_[1] uk_[3] ukurova_[1] ultra_[2] uluda&#number_[1] unbearable_[2] underline_[1] underwater_[2] unesco_[1] unplug_[3] unscramble_[1] upload_[5]

usa_[7] usage_[1] usb_[1] username_[1] utkan_[1] vacation_[13] vaccinate_[1] vaccinating_[1] vaccine_[2] vacuum_[6] vacuuming_[1] vacuums_[1] valentina_[1] valentino_[1] vancouver_[4] vanilla_[2] vegetable_[10] vegetables_[13] victor_[2] videos_[1] viewers_[1] virgin_[2] vocabulary_[3] voicemail_[4] voicemails_[1] volcanic_[3] volcano_[1] volleyball_[2] vwill_[1] waterfall_[1] waterfalls_[1] watson_[3] web_[9] webcam_[1] website_[11] weekdays_[9] wellknown_[1] wireless_[6] workbook_[1] wow_[1] wpuld_[1] wqvga_[1] wright_[5] www_[1] xaymaca_[1] xinhua_[1] yeah_[5] yearold_[1] yeast_[5] yell_[2] yoghurt_[3] yolk_[3] york_[1] yukon_[2] zealand_[5] zealanders_[3] zeynep_[3] zigana_[1] zorbing_[3] zucchinis_[1]

APPENDIX D

THE FREQUENCY OF THE WORDS IN THE COURSEBOOK

Word	Frequency	Relative Frequency
the	1749	6.25
and	831	2.97
a	605	2.16
you	580	2.07
I	514	1.84
to	514	1.84
is	490	1.75
in	489	1.75
your	348	1.24
it	347	1.24
do	346	1.24
of	342	1.22
are	306	1.09
can	254	0.91
with	239	0.85
what	237	0.85
about	230	0.82
my	215	0.77
for	207	0.74
on	201	0.72
unit	194	0.69
like	151	0.54
at	143	0.51
write	136	0.49
not	128	0.46
have	123	0.44
or	116	0.41
how	116	0.41
they	108	0.39
use	107	0.38
internet	99	0.35
we	96	0.34
am	94	0.34
read	94	0.34
an	94	0.34
he	87	0.31
would	83	0.30

Word	Frequency	Relative Frequency
make	83	0.30
going to	80	0.29
that	80	0.29
people	79	0.28
questions	76	0.27
there	71	0.25
when	70	0.25
below	69	0.25
but	68	0.24
activity	68	0.24
she	68	0.24
information	67	0.24
check	66	0.24
go	65	0.23
usually	65	0.23
school	64	0.23
sentences	63	0.23
text	63	0.23
lesson	60	0.21
this	59	0.21
s	59	0.21
be	58	0.21
answer	58	0.21
share	57	0.20
ask	55	0.20
simple	55	0.20
friends	54	0.19
does	53	0.19
help	53	0.19
example	52	0.19
very	52	0.19
photos	52	0.19
science	52	0.19
then	50	0.18
her	50	0.18
if	50	0.18

because	49	0.18	their	36	0.13
more	49	0.18	match	35	0.13
water	49	0.18	great	35	0.13
well	48	0.17	student	35	0.13
some	48	0.17	all	35	0.13
take	47	0.17	often	35	0.13
many	47	0.17	from	35	0.13
doing	46	0.16	online	35	0.13
them	46	0.16	paragraph	34	0.12
why	46	0.16	movies	33	0.12
understand	45	0.16	see	33	0.12
listen	44	0.16	box	33	0.12
put	44	0.16	little	33	0.12
please	44	0.16	study	32	0.11
one	43	0.15	favorite	32	0.11
short	43	0.15	leave	32	0.11
first	43	0.15	natural	32	0.11
so	42	0.15	think	31	0.11
world	42	0.15	sometimes	31	0.11
his	42	0.15	bike	31	0.11
will	42	0.15	by	31	0.11
work	41	0.15	communication	31	0.11
sports	41	0.15	these	30	0.11
hello	41	0.15	good	29	0.10
who	41	0.15	name	29	0.10
prefer	41	0.15	again	29	0.10
than	41	0.15	complete	29	0.10
home	40	0.14	class	29	0.10
me	39	0.14	time	29	0.10
after	39	0.14	boxes	29	0.10
play	38	0.14	has	29	0.10
follow	38	0.14	brother	29	0.10
was	38	0.14	dishes	29	0.10
most	38	0.14	hi	28	0.10
activities	37	0.13	yes	28	0.10
partner	37	0.13	correct	28	0.10
as	37	0.13	family	28	0.10
words	37	0.13	now	28	0.10
ok	37	0.13	also	28	0.10
chores	37	0.13	two	28	0.10
classmates	36	0.13	invitation	27	0.10
where	36	0.13	come	27	0.10

city	27	0.10	any	21	0.08
number	27	0.10	spoken	21	0.08
minutes	27	0.10	say	20	0.07
earthquake	27	0.10	conversation	20	0.07
us	26	0.09	three	20	0.07
always	26	0.09	get up	20	0.07
into	25	0.09	email	20	0.07
music	25	0.09	Bob	20	0.07
computer	25	0.09	connection	20	0.07
tourism	25	0.09	adventures	20	0.07
things	24	0.09	responsible	20	0.07
love	24	0.09	statements	20	0.07
friend	24	0.09	under	19	0.07
give	24	0.09	no	19	0.07
life	24	0.09	bye	19	0.07
parents	24	0.09	which	19	0.07
Dan	24	0.09	party	19	0.07
checklist	24	0.09	part	19	0.07
friendship	23	0.08	cooking	19	0.07
watch	23	0.08	add	19	0.07
movie	23	0.08	popular	19	0.07
messages	23	0.08	modem	19	0.07
message	23	0.08	sites	19	0.07
something	22	0.08	forces	19	0.07
visit	22	0.08	chart	18	0.06
fun	22	0.08	Helen	18	0.06
true	22	0.08	pairs	18	0.06
enjoy	22	0.08	find	18	0.06
teen	22	0.08	mean	18	0.06
recording	22	0.08	thanks	18	0.06
p.	22	0.08	texts	18	0.06
different	22	0.08	mother	18	0.06
points	22	0.08	homework	18	0.06
green	22	0.08	teens	18	0.06
did	22	0.08	such as	18	0.06
reasons	22	0.08	net	18	0.06
our	21	0.08	send	18	0.06
really	21	0.08	brochure	18	0.06
him	21	0.08	noun	18	0.06
daily	21	0.08	rafting	18	0.06
day	21	0.08	dangerous	18	0.06
could	21	0.08	practice	17	0.06

tomorrow	17	0.06	sorry	15	0.05
future	17	0.06	meet	15	0.05
place	17	0.06	letter	15	0.05
never	17	0.06	names	15	0.05
cycling	17	0.06	someone	15	0.05
father	17	0.06	numbers	15	0.05
afternoon	17	0.06	false	15	0.05
m	17	0.06	teacher	15	0.05
appropriate	17	0.06	before	15	0.05
ingredients	17	0.06	room	15	0.05
exchange	17	0.06	cover	15	0.05
note	17	0.06	express	15	0.05
look at	17	0.06	search	15	0.05
cable	17	0.06	language	15	0.05
its	17	0.06	mom	15	0.05
places	17	0.06	disasters	15	0.05
responsibilities	17	0.06	calling	14	0.05
evening	16	0.06	care	14	0.05
Frank	16	0.06	wash	14	0.05
anything	16	0.06	each	14	0.05
project	16	0.06	expressions	14	0.05
tell	16	0.06	best	14	0.05
walk	16	0.06	every day	14	0.05
house	16	0.06	get	14	0.05
breakfast	16	0.06	Kate	14	0.05
tv	16	0.06	listening	14	0.05
Sam	16	0.06	four	14	0.05
answers	16	0.06	process	14	0.05
talk	16	0.06	describe	14	0.05
cook	16	0.06	while	14	0.05
pan	16	0.06	ideas	14	0.05
salt	16	0.06	around	14	0.05
using	16	0.06	parkour	14	0.05
phrases	16	0.06	running	14	0.05
extreme sports	16	0.06	tourist attractions	14	0.05
pocket	16	0.06	your own	14	0.05
historic	16	0.06	turn off	14	0.05
important	16	0.06	dossier	14	0.05
should	16	0.06	kind	13	0.05
clean	16	0.06	shopping	13	0.05
just	15	0.05	week	13	0.05
ones	15	0.05	above	13	0.05

Saturday	13	0.05	hot	12	0.04
refuse	13	0.05	call	12	0.04
dictionary	13	0.05	phone	12	0.04
a little	13	0.05	at the moment	12	0.04
teeth	13	0.05	interaction	12	0.04
habits	13	0.05	vacation	12	0.04
Turkey	13	0.05	racing	12	0.04
take out	13	0.05	reading	12	0.04
vegetables	13	0.05	other	11	0.04
dish	13	0.05	excuses	11	0.04
speak	13	0.05	show	11	0.04
between	13	0.05	decide	11	0.04
sister	13	0.05	examples	11	0.04
hang	13	0.05	exciting	11	0.04
list	13	0.05	cycle	11	0.04
poster	13	0.05	m.	11	0.04
during	13	0.05	preferences	11	0.04
writing	13	0.05	keep	11	0.04
earth	13	0.05	bowl	11	0.04
having	12	0.04	flour	11	0.04
together	12	0.04	large	11	0.04
point	12	0.04	rice	11	0.04
Mike	12	0.04	need	11	0.04
over	12	0.04	million	11	0.04
plans	12	0.04	new	11	0.04
accept	12	0.04	website	11	0.04
join	12	0.04	I mean	11	0.04
b	12	0.04	main	11	0.04
hope	12	0.04	today	11	0.04
scan	12	0.04	were	11	0.04
likes	12	0.04	person	11	0.04
do you think	12	0.04	museum	11	0.04
let	12	0.04	dna	11	0.04
interesting	12	0.04	Onda	11	0.04
early	12	0.04	late	10	0.04
morning	12	0.04	stay	10	0.04
dinner	12	0.04	weekend	10	0.04
design	12	0.04	compare	10	0.04
bread	12	0.04	talking	10	0.04
butter	12	0.04	imagine	10	0.04
finally	12	0.04	start	10	0.04
here	12	0.04	emails	10	0.04

dislikes	10	0.04		invitations	9	0.03
food	10	0.04		tonight	9	0.03
try	10	0.04		soccer	9	0.03
too	10	0.04		Sally	9	0.03
problem	10	0.04		state	9	0.03
news	10	0.04		you know	9	0.03
member	10	0.04		others	9	0.03
eat	10	0.04		may	9	0.03
bad	10	0.04		Donna	9	0.03
vegetable	10	0.04		idea	9	0.03
pizza	10	0.04		parts	9	0.03
oil	10	0.04		off	9	0.03
glass	10	0.04		yourself	9	0.03
hours	10	0.04		routines	9	0.03
social	10	0.04		live	9	0.03
clarification	10	0.04		milk	9	0.03
Bill	10	0.04		weekdays	9	0.03
uses	10	0.04		cheese	9	0.03
kayaking	10	0.04		pop	9	0.03
challenging	10	0.04		mostly	9	0.03
fascinating	10	0.04		prepare	9	0.03
entertaining	10	0.04		fry	9	0.03
buildings	10	0.04		bake	9	0.03
mosque	10	0.04		boil	9	0.03
weather	10	0.04		next	9	0.03
tourist	10	0.04		heat	9	0.03
word	10	0.04		until	9	0.03
floor	10	0.04		display	9	0.03
poem	10	0.04		want	9	0.03
bulb	10	0.04		fast	9	0.03
base	10	0.04		teenagers	9	0.03
assess	10	0.04		web	9	0.03
progress	10	0.04		connect	9	0.03
checking	10	0.04		pc	9	0.03
go back	10	0.04		through	9	0.03
item	10	0.04		screen	9	0.03
filled in	10	0.04		identify	9	0.03
filled	10	0.04		diving	9	0.03
cut out	10	0.04		skateboarding	9	0.03
production	10	0.04		parachute	9	0.03
watching	9	0.03		adventure	9	0.03
sci-fi	9	0.03		real	9	0.03

size	9	0.03	okay	8	0.03
visiting	9	0.03	plan	8	0.03
modern	9	0.03	repeat	8	0.03
blue	9	0.03	later	8	0.03
table	9	0.03	playing	8	0.03
rules	9	0.03	technology	8	0.03
invention	9	0.03	approximately	8	0.03
invented	9	0.03	hard	8	0.03
scientific	9	0.03	much	8	0.03
bring	9	0.03	users	8	0.03
lessons	9	0.03	at least	8	0.03
computer games	8	0.03	gliding	8	0.03
Sunday	8	0.03	white	8	0.03
going	8	0.03	structure	8	0.03
bus	8	0.03	designed	8	0.03
definitions	8	0.03	climate	8	0.03
enough	8	0.03	Jamaica	8	0.03
gale	8	0.03	garbage	8	0.03
years	8	0.03	dust	8	0.03
snow	8	0.03	must	8	0.03
outside	8	0.03	feelings	8	0.03
drive	8	0.03	said	8	0.03
shower	8	0.03	aircraft	8	0.03
brush	8	0.03	working	8	0.03
eggs	8	0.03	interested	8	0.03
band	8	0.03	tsunami	8	0.03
voice	8	0.03	landslide	8	0.03
sport	8	0.03	save	8	0.03
groups	8	0.03	Australia	8	0.03
cut	8	0.03	networking	8	0.03
toast	8	0.03	Mardin	8	0.03
jam	8	0.03	thrillers	7	0.03
recipe	8	0.03	busy	7	0.03
pepper	8	0.03	have to	7	0.03
onion	8	0.03	making	7	0.03
set	8	0.03	question	7	0.03
sugar	8	0.03	nature	7	0.03
stir	8	0.03	excuse	7	0.03
dough	8	0.03	example example	7	0.03
hour	8	0.03	brothers	7	0.03
eye-catching	8	0.03	spend	7	0.03
available	8	0.03	lunch	7	0.03

starts	7	0.03	needed	7	0.03
choose	7	0.03	poly	7	0.03
make it	7	0.03	regular	7	0.03
easy	7	0.03	compensation	7	0.03
tablespoons	7	0.03	strategies	7	0.03
pour	7	0.03	drink	6	0.02
mixture	7	0.03	grandparents	6	0.02
Turkish	7	0.03	change	6	0.02
means	7	0.03	invite	6	0.02
five	7	0.03	oh	6	0.02
conversations	7	0.03	sure	6	0.02
Curtis	7	0.03	organizing	6	0.02
hear	7	0.03	same	6	0.02
lots	7	0.03	know	6	0.02
Marc	7	0.03	called	6	0.02
personal	7	0.03	boring	6	0.02
Jason	7	0.03	sea	6	0.02
thing	7	0.03	ride	6	0.02
piece of paper	7	0.03	hardly ever	6	0.02
canoeing	7	0.03	clock	6	0.02
comparisons	7	0.03	weekends	6	0.02
easier	7	0.03	photo	6	0.02
adrenalin	7	0.03	a.	6	0.02
rather	7	0.03	rock	6	0.02
move	7	0.03	once or twice	6	0.02
year	7	0.03	interview	6	0.02
special	7	0.03	related	6	0.02
Tarsus	7	0.03	grill	6	0.02
capital	7	0.03	oven	6	0.02
island	7	0.03	cabbage	6	0.02
destination	7	0.03	carrot	6	0.02
dishwasher	7	0.03	dry	6	0.02
housework	7	0.03	wait	6	0.02
makes	7	0.03	pasta	6	0.02
obligations	7	0.03	countries	6	0.02
inventions	7	0.03	sliced	6	0.02
include	7	0.03	big	6	0.02
brushing	7	0.03	traditional	6	0.02
last	7	0.03	self-assessment	6	0.02
fourteen	7	0.03	in touch with	6	0.02
Danny	7	0.03	face to face	6	0.02
tennis	7	0.03	dad	6	0.02

come to	6	0.02	space	6	0.02
line	6	0.02	Argentina	6	0.02
Linda	6	0.02	caused	6	0.02
respond	6	0.02	tornado	6	0.02
soon	6	0.02	disaster	6	0.02
nearly	6	0.02	kilometres	6	0.02
according to	6	0.02	hit	6	0.02
etc.	6	0.02	waste	6	0.02
log	6	0.02	causes	6	0.02
English	6	0.02	recordings	6	0.02
tablet	6	0.02	jack	6	0.02
wireless	6	0.02	small	6	0.02
internet addicts	6	0.02	had	6	0.02
games	6	0.02	made	6	0.02
caving	6	0.02	comedies	5	0.02
I think	6	0.02	come over	5	0.02
less	6	0.02	not really	5	0.02
free time	6	0.02	yeah	5	0.02
motorbike	6	0.02	night	5	0.02
bikes	6	0.02	soccer match	5	0.02
USA	6	0.02	circle	5	0.02
long	6	0.02	lines	5	0.02
tall	6	0.02	birthday	5	0.02
mountains	6	0.02	bicycle	5	0.02
province	6	0.02	walking	5	0.02
passages	6	0.02	end	5	0.02
shelves	6	0.02	go with	5	0.02
washing	6	0.02	ready	5	0.02
vacuum	6	0.02	guess	5	0.02
feed	6	0.02	address	5	0.02
bathroom	6	0.02	events	5	0.02
Monday	6	0.02	common	5	0.02
desk	6	0.02	stranger	5	0.02
plants	6	0.02	lives	5	0.02
building	6	0.02	become	5	0.02
medicine	6	0.02	feel	5	0.02
strong	6	0.02	offers	5	0.02
electric light	6	0.02	hiking	5	0.02
powered	6	0.02	mountain	5	0.02
developed	6	0.02	near	5	0.02
achievements	6	0.02	old	5	0.02
separate	6	0.02	go to bed	5	0.02

Italy	5	0.02	motor racing	5	0.02
students	5	0.02	quickly	5	0.02
clothes	5	0.02	brochures	5	0.02
tickets	5	0.02	0stanbul	5	0.02
cd	5	0.02	location	5	0.02
draw	5	0.02	stop	5	0.02
snack	5	0.02	various	5	0.02
slice	5	0.02	experience	5	0.02
describing	5	0.02	cuisine	5	0.02
French	5	0.02	preference	5	0.02
tomato	5	0.02	chore	5	0.02
fried	5	0.02	bed	5	0.02
top	5	0.02	empty	5	0.02
delicious	5	0.02	August	5	0.02
yeast	5	0.02	newspaper	5	0.02
tablespoon	5	0.02	Wednesday	5	0.02
warm	5	0.02	children	5	0.02
roll	5	0.02	windows	5	0.02
beef	5	0.02	classroom	5	0.02
hands	5	0.02	shoes	5	0.02
serve	5	0.02	bath	5	0.02
baking powder	5	0.02	fold	5	0.02
half	5	0.02	structures	5	0.02
phone call	5	0.02	take care of	5	0.02
Martin	5	0.02	necessary	5	0.02
outline	5	0.02	discovery	5	0.02
inform	5	0.02	inventor	5	0.02
stores	5	0.02	took	5	0.02
mobile phones	5	0.02	phonograph	5	0.02
even	5	0.02	mechanical	5	0.02
books	5	0.02	telescope	5	0.02
because of	5	0.02	Wright	5	0.02
download	5	0.02	built	5	0.02
upload	5	0.02	few	5	0.02
access	5	0.02	destroy	5	0.02
everywhere	5	0.02	universe	5	0.02
site	5	0.02	possible	5	0.02
mobile	5	0.02	specific	5	0.02
problems	5	0.02	chemist	5	0.02
homes	5	0.02	global	5	0.02
month	5	0.02	wave	5	0.02
bungee jumping	5	0.02	ground	5	0.02

rain	5	0.02		tournament	4	0.01
damage	5	0.02		race	4	0.01
drought	5	0.02		picnic	4	0.01
avalanche	5	0.02		greet	4	0.01
per	5	0.02		Tuesday	4	0.01
died	5	0.02		order	4	0.01
injured	5	0.02		count	4	0.01
missing	5	0.02		forever	4	0.01
fall	5	0.02		act out	4	0.01
saving	5	0.02		self-assessment	4	0.01
suggestions	5	0.02		another	4	0.01
reduce	5	0.02		better	4	0.01
health	5	0.02		apologizing	4	0.01
fluorescent	5	0.02		adjectives	4	0.01
recycle	5	0.02		biking	4	0.01
predictions	5	0.02		tubing	4	0.01
forest	5	0.02		trendy	4	0.01
Nancy	5	0.02		neighbourhood	4	0.01
jumping	5	0.02		goes	4	0.01
speed	5	0.02		mudguard	4	0.01
architecture	5	0.02		handlebar	4	0.01
meters	5	0.02		frame	4	0.01
local	5	0.02		brake lever	4	0.01
energy	5	0.02		pedal	4	0.01
ally	5	0.02		tire	4	0.01
Cansu Ellen	5	0.02		seat	4	0.01
Gina	5	0.02		away from	4	0.01
New Zealand	5	0.02		Sundays	4	0.01
delete	5	0.02		o	4	0.01
spam	5	0.02		honey	4	0.01
collapse	5	0.02		frequency	4	0.01
go for	4	0.01		coffee	4	0.01
posters	4	0.01		Italian	4	0.01
types	4	0.01		routine	4	0.01
romances	4	0.01		times	4	0.01
comedy	4	0.01		notes	4	0.01
accepting	4	0.01		singer	4	0.01
cousin	4	0.01		jazz	4	0.01
feeling	4	0.01		hang out	4	0.01
uncle	4	0.01		terrific	4	0.01
accepts	4	0.01		impressive	4	0.01
refuses	4	0.01		loud	4	0.01

fruit	4	0.01	teachers	4	0.01
score	4	0.01	messaging	4	0.01
a lot	4	0.01	languages	4	0.01
forget	4	0.01	nice	4	0.01
cool	4	0.01	account	4	0.01
create	4	0.01	register	4	0.01
pictures	4	0.01	engine	4	0.01
following	4	0.01	browser	4	0.01
fish	4	0.01	copy	4	0.01
chicken	4	0.01	password	4	0.01
carrots	4	0.01	desktop	4	0.01
spread	4	0.01	telephone	4	0.01
slices	4	0.01	sign	4	0.01
frying	4	0.01	clear	4	0.01
meanings	4	0.01	asking	4	0.01
mixing	4	0.01	offline	4	0.01
loaf	4	0.01	being	4	0.01
teaspoon	4	0.01	internet addict	4	0.01
surface	4	0.01	only	4	0.01
shape	4	0.01	chat	4	0.01
cold	4	0.01	find out	4	0.01
orzo	4	0.01	several times	4	0.01
low	4	0.01	topic	4	0.01
brown	4	0.01	sky	4	0.01
aye	4	0.01	right	4	0.01
national	4	0.01	Joe	4	0.01
meat	4	0.01	seeker	4	0.01
meal	4	0.01	against	4	0.01
description	4	0.01	in turn	4	0.01
mix	4	0.01	scenery	4	0.01
step	4	0.01	explore	4	0.01
voicemail	4	0.01	fantastic	4	0.01
Rachael	4	0.01	tourists	4	0.01
Tina	4	0.01	pyramid	4	0.01
Peter	4	0.01	Egyptian	4	0.01
stadium	4	0.01	beautiful	4	0.01
Robert	4	0.01	truly	4	0.01
pretend	4	0.01	accommodation	4	0.01
abbreviations	4	0.01	largest	4	0.01
service	4	0.01	every	4	0.01
billion	4	0.01	topics	4	0.01
buy	4	0.01	paper	4	0.01

grocery	4	0.01	wind	4	0.01
laundry	4	0.01	flood	4	0.01
taken	4	0.01	Japanese	4	0.01
diary	4	0.01	Richter	4	0.01
Jeremy	4	0.01	scale	4	0.01
happy	4	0.01	result	4	0.01
garden	4	0.01	shake	4	0.01
cleaning	4	0.01	massive	4	0.01
Denis	4	0.01	left	4	0.01
Friday	4	0.01	elevator	4	0.01
mop	4	0.01	debris	4	0.01
given	4	0.01	sturdy	4	0.01
dog	4	0.01	head	4	0.01
pick up	4	0.01	arms	4	0.01
cooks	4	0.01	tap	4	0.01
inventors	4	0.01	protect	4	0.01
discover	4	0.01	planet	4	0.01
discovered	4	0.01	bin	4	0.01
Nobel prize	4	0.01	light bulbs	4	0.01
electricity	4	0.01	plastic	4	0.01
motion picture	4	0.01	reuse	4	0.01
camera	4	0.01	support	4	0.01
practical	4	0.01	seven	4	0.01
powerful	4	0.01	popular in	4	0.01
paragraphs	4	0.01	started	4	0.01
fly	4	0.01	went	4	0.01
Nanomedicine	4	0.01	run	4	0.01
experiments	4	0.01	center	4	0.01
computers	4	0.01	hey	4	0.01
down	4	0.01	sad	4	0.01
putting	4	0.01	Rob	4	0.01
chemical	4	0.01	roly	4	0.01
test-tube	4	0.01	desserts	4	0.01
area	4	0.01	choices	4	0.01
melts	4	0.01	postcard	4	0.01
warming	4	0.01	Spanish	4	0.01
high	4	0.01	keyboard	4	0.01
park	4	0.01	rock-climbing	4	0.01
involve	4	0.01	helmet	4	0.01
descriptions	4	0.01	first-aid	4	0.01
violent	4	0.01	Brian	4	0.01
movement	4	0.01	Sandra	4	0.01

tower	4	0.01	verb	3	0.01
wheel	4	0.01	forms	3	0.01
wheels	4	0.01	members	3	0.01
regularly	4	0.01	pancakes	3	0.01
recognize	4	0.01	teenager	3	0.01
past	4	0.01	orange juice	3	0.01
caf	3	0.01	rest	3	0.01
goodbye	3	0.01	beach	3	0.01
romance	3	0.01	wear	3	0.01
refusing	3	0.01	fashion	3	0.01
coming	3	0.01	classical	3	0.01
go around	3	0.01	songs	3	0.01
physically	3	0.01	energetic	3	0.01
active	3	0.01	book	3	0.01
Zeynep	3	0.01	exercise	3	0.01
aunt	3	0.01	headphones	3	0.01
exam	3	0.01	performer	3	0.01
test	3	0.01	Clare	3	0.01
Ashley	3	0.01	guitar	3	0.01
barbecue	3	0.01	eggplants	3	0.01
role-play	3	0.01	roast	3	0.01
sender	3	0.01	explain	3	0.01
read out	3	0.01	snacks	3	0.01
say goodbye	3	0.01	peel	3	0.01
difficult	3	0.01	chop	3	0.01
date	3	0.01	dice	3	0.01
event	3	0.01	close	3	0.01
back up	3	0.01	chopped	3	0.01
get on	3	0.01	microwave	3	0.01
turns	3	0.01	piece	3	0.01
happen	3	0.01	rolling	3	0.01
supposed	3	0.01	pin	3	0.01
serving	3	0.01	knead	3	0.01
drinks	3	0.01	glasses	3	0.01
bossaball	3	0.01	recipes	3	0.01
zorbing	3	0.01	melt	3	0.01
snowshoeing	3	0.01	rinse	3	0.01
unusual	3	0.01	turn down	3	0.01
hanging out	3	0.01	Mehmet	3	0.01
riverside	3	0.01	pilaf	3	0.01
fixing	3	0.01	afghan	3	0.01
chain	3	0.01	lamb	3	0.01

term	3	0.01	expert	3	0.01
colcannon	3	0.01	performing	3	0.01
contain	3	0.01	pocket-bike	3	0.01
booklet	3	0.01	brackets	3	0.01
yoghurt	3	0.01	tried	3	0.01
Steve	3	0.01	concept	3	0.01
contact	3	0.01	mini	3	0.01
caller	3	0.01	standard	3	0.01
receiver	3	0.01	consider	3	0.01
answering	3	0.01	trip	3	0.01
moment	3	0.01	decided	3	0.01
speaking	3	0.01	Antalya	3	0.01
sympathy	3	0.01	narratives	3	0.01
board games	3	0.01	suitable	3	0.01
free	3	0.01	cities	3	0.01
theater	3	0.01	Hagia Sophia	3	0.01
cheap	3	0.01	Eiffel tower	3	0.01
young	3	0.01	London	3	0.01
generally	3	0.01	UK	3	0.01
for example	3	0.01	ancient	3	0.01
believe	3	0.01	Niagara Falls	3	0.01
mock	3	0.01	construction	3	0.01
concerns	3	0.01	expressing	3	0.01
clarify	3	0.01	giving	3	0.01
myself	3	0.01	history	3	0.01
Saturday afternoon	3	0.01	sightseeing	3	0.01
attachment	3	0.01	recreational	3	0.01
out	3	0.01	km	3	0.01
memory	3	0.01	civilizations	3	0.01
post	3	0.01	Mediterranean	3	0.01
software	3	0.01	region	3	0.01
trying	3	0.01	winters	3	0.01
open	3	0.01	grand	3	0.01
follow up	3	0.01	Caribbean	3	0.01
connected	3	0.01	tropical	3	0.01
all the time	3	0.01	beaches	3	0.01
title	3	0.01	waters	3	0.01
maybe	3	0.01	third	3	0.01
connectors	3	0.01	huge	3	0.01
definition	3	0.01	coconut	3	0.01
seekers	3	0.01	tomb	3	0.01
danger	3	0.01	recorded	3	0.01

grocery shopping	3	0.01	thesis	3	0.01
tidy up	3	0.01	glaciers	3	0.01
meals	3	0.01	glacial	3	0.01
responsibility	3	0.01	knowledge	3	0.01
dusting	3	0.01	environmental	3	0.01
finished	3	0.01	shows	3	0.01
car	3	0.01	located	3	0.01
drinking	3	0.01	spectacular	3	0.01
floors	3	0.01	unique	3	0.01
Betsy	3	0.01	extremely	3	0.01
excited	3	0.01	amount	3	0.01
hate	3	0.01	period	3	0.01
Thursday	3	0.01	hurricane	3	0.01
kitchen	3	0.01	magnitude	3	0.01
dirty	3	0.01	waves	3	0.01
put away	3	0.01	up	3	0.01
as soon as	3	0.01	volcanic	3	0.01
discoveries	3	0.01	Mexico	3	0.01
discoverer	3	0.01	shook	3	0.01
invent	3	0.01	injuries	3	0.01
material	3	0.01	trapped	3	0.01
compass	3	0.01	crawl	3	0.01
screw	3	0.01	walls	3	0.01
Thomas Edison	3	0.01	immediately	3	0.01
develop	3	0.01	shaking	3	0.01
successfully	3	0.01	stops	3	0.01
showed	3	0.01	hold	3	0.01
man	3	0.01	drop	3	0.01
including	3	0.01	water shortages	3	0.01
among	3	0.01	too much	3	0.01
ten	3	0.01	wasting	3	0.01
computing	3	0.01	shorter	3	0.01
details	3	0.01	baths	3	0.01
cancer cells	3	0.01	taps	3	0.01
storing	3	0.01	greenhouse	3	0.01
unlimited	3	0.01	gases	3	0.01
prototype	3	0.01	calories	3	0.01
capable	3	0.01	public	3	0.01
nanotechnology	3	0.01	transportation	3	0.01
substance	3	0.01	pollution	3	0.01
scientists	3	0.01	soda	3	0.01
major	3	0.01	compact	3	0.01

rarely	3	0.01	jam Roly	3	0.01
unplug	3	0.01	Australians	3	0.01
light	3	0.01	Zealanders	3	0.01
metal	3	0.01	summer	3	0.01
double	3	0.01	both	3	0.01
scores	3	0.01	letters	3	0.01
gather	3	0.01	receive	3	0.01
flooding	3	0.01	communicate	3	0.01
heavy	3	0.01	asked	3	0.01
break	3	0.01	safety	3	0.01
heard	3	0.01	research	3	0.01
Tim	3	0.01	jumpers	3	0.01
Bruce	3	0.01	jump	3	0.01
Ahmet	3	0.01	broken	3	0.01
travelling	3	0.01	kit	3	0.01
Dave	3	0.01	prize	3	0.01
stove	3	0.01	learn	3	0.01
tomato sauce	3	0.01	monastery	3	0.01
com	3	0.01	China	3	0.01
designer	3	0.01	's	3	0.01
wide	3	0.01	helps	3	0.01
motorbikes	3	0.01	my own	3	0.01
began	3	0.01	scientist	3	0.01
famous	3	0.01	brought	3	0.01
Emperor	3	0.01	fair	3	0.01
border	3	0.01	stable	3	0.01
clock tower	3	0.01	rider	3	0.01
tallest	3	0.01	adjust	3	0.01
historical	3	0.01	combination	3	0.01
reports	3	0.01	comparing	3	0.01
authorities	3	0.01	present	3	0.01
safe	3	0.01	actions	3	0.01
evenings	3	0.01	attitudes	3	0.01
crossword puzzle	3	0.01	intercultural	3	0.01
Henry	3	0.01	awareness	3	0.01
six	3	0.01	discussion	3	0.01
dear	3	0.01	highest	3	0.01
eating	3	0.01	vocabulary	3	0.01
Vancouver CANADA	3	0.01	form	3	0.01
cm	3	0.01	store	3	0.01
g	3	0.01	exhibition	3	0.01
egg yolk	3	0.01	precious	3	0.01

Tarsus	3	0.01	everything	2	0.01
walks	2	0.01	in response to	2	0.01
plays	2	0.01	attending	2	0.01
eat out	2	0.01	inviting	2	0.01
shopping malls	2	0.01	Fred	2	0.01
sit	2	0.01	black	2	0.01
chit-chat	2	0.01	silver	2	0.01
movie types	2	0.01	mudguards	2	0.01
Frank Mike Frank	2	0.01	brake gear	2	0.01
Frank Mike Mike	2	0.01	high school	2	0.01
not at all	2	0.01	far	2	0.01
thriller	2	0.01	noon	2	0.01
responses	2	0.01	gets up	2	0.01
go out	2	0.01	Daniela	2	0.01
5	2	0.01	gets	2	0.01
further	2	0.01	the afternoon after	2	0.01
out-of-town	2	0.01	of the day	2	0.01
relatives	2	0.01	for a while	2	0.01
Pete	2	0.01	sleep	2	0.01
volleyball	2	0.01	fourteen-year-old	2	0.01
shopping mall	2	0.01	Daniela Elmo	2	0.01
tennis match	2	0.01	tells	2	0.01
friendship work	2	0.01	alone	2	0.01
cards	2	0.01	casual	2	0.01
reply	2	0.01	Kate Sam	2	0.01
attend to	2	0.01	Taylor Swift Sam	2	0.01
sounds	2	0.01	unbearable	2	0.01
Amanda	2	0.01	quiz	2	0.01
buddies	2	0.01	color	2	0.01
each other	2	0.01	twice	2	0.01
come into	2	0.01	once-a-month	2	0.01
sisters	2	0.01	fast-food	2	0.01
lie	2	0.01	junk	2	0.01
in common	2	0.01	magazine	2	0.01
happens	2	0.01	quite	2	0.01
similar	2	0.01	lifestyle	2	0.01
dialog	2	0.01	instruments	2	0.01
enjoy yourself	2	0.01	flute	2	0.01
cup of tea	2	0.01	organ	2	0.01
animation	2	0.01	song	2	0.01
animated	2	0.01	cooking cooking	2	0.01
drama	2	0.01	roast steam	2	0.01

mushrooms	2	0.01	invites	2	0.01
onions	2	0.01	Rachael Curtis	2	0.01
steam	2	0.01	afraid	2	0.01
toaster	2	0.01	gone out	2	0.01
turn on	2	0.01	Mr. Bright	2	0.01
red	2	0.01	Curtis Rachael	2	0.01
in half	2	0.01	Tina speaking	2	0.01
steps	2	0.01	Peter Tina	2	0.01
sequencing	2	0.01	hang on a minute	2	0.01
floured	2	0.01	Martin Peter Peter	2	0.01
ball	2	0.01	west high	2	0.01
remove	2	0.01	Yukon	2	0.01
discuss	2	0.01	stadium Martin	2	0.01
write down	2	0.01	start Peter	2	0.01
saucepan	2	0.01	Martin Good	2	0.01
cooker	2	0.01	right now	2	0.01
stirring	2	0.01	hold on a moment	2	0.01
continuously	2	0.01	Matt speaking	2	0.01
absorbs	2	0.01	Mrs. Adams	2	0.01
cool down	2	0.01	Mrs. Peterson	2	0.01
lasagne	2	0.01	pardon	2	0.01
consists	2	0.01	Jill Sanders	2	0.01
lentils	2	0.01	Mr. Gino	2	0.01
nuts	2	0.01	Albert Parker	2	0.01
pistachios	2	0.01	Eric	2	0.01
almonds	2	0.01	concern	2	0.01
middle	2	0.01	meet up	2	0.01
Beshbarmak	2	0.01	barbeque	2	0.01
noodles	2	0.01	accessory	2	0.01
Irish	2	0.01	b4n	2	0.01
mashed potatoes	2	0.01	secret	2	0.01
regional	2	0.01	became	2	0.01
variations	2	0.01	schools	2	0.01
recipe booklet	2	0.01	effects	2	0.01
unknown	2	0.01	changing	2	0.01
berries	2	0.01	communicating	2	0.01
muffin	2	0.01	face-to-face	2	0.01
fresh	2	0.01	hand	2	0.01
ways	2	0.01	record	2	0.01
fax	2	0.01	hang up	2	0.01
networks	2	0.01	agree	2	0.01
formal	2	0.01	comment	2	0.01

confirm	2	0.01	objects	2	0.01
log in	2	0.01	skillful	2	0.01
disconnect	2	0.01	movements	2	0.01
computer program	2	0.01	country	2	0.01
entering	2	0.01	begin	2	0.01
opinion	2	0.01	manufacturers	2	0.01
make friends	2	0.01	producing	2	0.01
chats	2	0.01	extreme sport	2	0.01
notebook	2	0.01	fans	2	0.01
magazines	2	0.01	riding	2	0.01
keeps	2	0.01	hot air balloon	2	0.01
disk	2	0.01	attractions	2	0.01
processor	2	0.01	statue	2	0.01
Dan speaking	2	0.01	liberty	2	0.01
Bob Dan	2	0.01	Taj Mahal	2	0.01
How are things	2	0.01	architectural style	2	0.01
Bob Bob Dan	2	0.01	every year	2	0.01
That is it	2	0.01	Big Ben	2	0.01
Thanks a lot	2	0.01	workers	2	0.01
never mind	2	0.01	build	2	0.01
solve	2	0.01	type	2	0.01
sing	2	0.01	seaside	2	0.01
smartphone	2	0.01	multicultural	2	0.01
addicts	2	0.01	central	2	0.01
families	2	0.01	Mediterranean Sea	2	0.01
counselling	2	0.01	roman	2	0.01
book fair	2	0.01	scene	2	0.01
questionnaire	2	0.01	plain	2	0.01
follow-up	2	0.01	typical	2	0.01
given in	2	0.01	summers	2	0.01
harder	2	0.01	chilly	2	0.01
c	2	0.01	damp	2	0.01
and other	2	0.01	gate	2	0.01
risk	2	0.01	bridge	2	0.01
raft	2	0.01	church	2	0.01
rivers	2	0.01	tracking	2	0.01
freedom	2	0.01	tiring	2	0.01
themselves	2	0.01	kinds	2	0.01
air	2	0.01	all-inclusive	2	0.01
mysterious	2	0.01	resort	2	0.01
scuba diving	2	0.01	bed and breakfast	2	0.01
reason	2	0.01	official	2	0.01

population	2	0.01	Leo	2	0.01
currency	2	0.01	has to	2	0.01
dollar	2	0.01	sweep	2	0.01
friendly	2	0.01	lists	2	0.01
majestic	2	0.01	directions	2	0.01
sandy	2	0.01	every morning	2	0.01
land	2	0.01	tidy	2	0.01
mild	2	0.01	raise	2	0.01
temperature	2	0.01	math	2	0.01
banana trees	2	0.01	pants	2	0.01
palms	2	0.01	socks	2	0.01
flowers	2	0.01	fishes	2	0.01
looking for	2	0.01	blocks	2	0.01
memorable	2	0.01	carpets	2	0.01
exploring	2	0.01	shirts	2	0.01
Bursa	2	0.01	potted	2	0.01
winter	2	0.01	done	2	0.01
ski resort	2	0.01	mopping	2	0.01
definitely	2	0.01	furniture	2	0.01
dessert	2	0.01	talks	2	0.01
budget	2	0.01	respect	2	0.01
destinations	2	0.01	discoverers	2	0.01
tourist attraction	2	0.01	kevlar	2	0.01
peer	2	0.01	synthetic	2	0.01
assessment	2	0.01	steel	2	0.01
gives	2	0.01	improved	2	0.01
effectively	2	0.01	flew	2	0.01
grammar	2	0.01	twentieth century	2	0.01
spelling	2	0.01	surf	2	0.01
punctuation	2	0.01	devices	2	0.01
capitalization	2	0.01	microchip	2	0.01
household chores	2	0.01	diagnose	2	0.01
clean up	2	0.01	aids	2	0.01
ironing	2	0.01	within	2	0.01
load	2	0.01	tested	2	0.01
entries	2	0.01	outstanding	2	0.01
getting	2	0.01	innovation	2	0.01
tidying up	2	0.01	award	2	0.01
bedrooms	2	0.01	supporting	2	0.01
tasks	2	0.01	innovations	2	0.01
yesterday	2	0.01	limited	2	0.01
tired	2	0.01	planning	2	0.01

perform	2	0.01	ice	2	0.01
surgeries	2	0.01	sudden	2	0.01
genetic	2	0.01	mass	2	0.01
prototypes	2	0.01	take place	2	0.01
nanomolecules	2	0.01	happened	2	0.01
ultra-delicate	2	0.01	affect	2	0.01
bacteria	2	0.01	December	2	0.01
biggest	2	0.01	coast	2	0.01
breakthrough	2	0.01	the ocean	2	0.01
vaccine	2	0.01	Africa	2	0.01
science work	2	0.01	giant	2	0.01
remember	2	0.01	still	2	0.01
student B	2	0.01	remain	2	0.01
girl	2	0.01	homeless	2	0.01
diseases	2	0.01	underwater	2	0.01
Bunsen	2	0.01	meteor	2	0.01
burner	2	0.01	onto	2	0.01
taking	2	0.01	highlights	2	0.01
technician	2	0.01	Pacific coast	2	0.01
Katya	2	0.01	breaking news	2	0.01
studying	2	0.01	measured	2	0.01
ecosystems	2	0.01	buried	2	0.01
improve	2	0.01	suburbs	2	0.01
articles	2	0.01	Oklahoma City	2	0.01
engineering	2	0.01	yell	2	0.01
glacier	2	0.01	inside	2	0.01
discussions	2	0.01	panic	2	0.01
achievement	2	0.01	sprinkler	2	0.01
astonishing	2	0.01	cloth	2	0.01
welcome	2	0.01	noise	2	0.01
interactive	2	0.01	once	2	0.01
volunteer	2	0.01	quake	2	0.01
3d	2	0.01	electric	2	0.01
allow	2	0.01	wall	2	0.01
cosmos	2	0.01	water shortage	2	0.01
culture	2	0.01	pay attention	2	0.01
astronomy	2	0.01	every time	2	0.01
way	2	0.01	doing good	2	0.01
wonder	2	0.01	anywhere	2	0.01
circular	2	0.01	burning	2	0.01
cone	2	0.01	improving	2	0.01
moves	2	0.01	longer	2	0.01

two thirds	2	0.01	wonders	2	0.01
monitor	2	0.01	ordered	2	0.01
air conditioner	2	0.01	amazing	2	0.01
plastic bags	2	0.01	ottoman	2	0.01
shops	2	0.01	architect	2	0.01
print	2	0.01	man-made	2	0.01
sided	2	0.01	palace	2	0.01
pale	2	0.01	check-out	2	0.01
bright	2	0.01	beds	2	0.01
opinions	2	0.01	fifteen	2	0.01
without	2	0.01	cakes	2	0.01
crops	2	0.01	once a week	2	0.01
die	2	0.01	second	2	0.01
occur	2	0.01	patents	2	0.01
desert	2	0.01	products	2	0.01
floods	2	0.01	Archimedes	2	0.01
overflow	2	0.01	Charles Babbage	2	0.01
centimeters	2	0.01	telephone number	2	0.01
flash floods	2	0.01	rescue	2	0.01
Suzy	2	0.01	nine	2	0.01
Abby	2	0.01	team	2	0.01
gave	2	0.01	gym	2	0.01
ray	2	0.01	cause	2	0.01
Kevin	2	0.01	whenever	2	0.01
theatres	2	0.01	call on	2	0.01
especially	2	0.01	basketball	2	0.01
Jane	2	0.01	verbs	2	0.01
Saturday morning	2	0.01	mother work	2	0.01
brake	2	0.01	afternoons	2	0.01
fix	2	0.01	train	2	0.01
corns	2	0.01	meaningful	2	0.01
mixed	2	0.01	actually	2	0.01
Internet Jason	2	0.01	pen pal	2	0.01
jasonvill	2	0.01	convey	2	0.01
I guess	2	0.01	swimming	2	0.01
almost	2	0.01	seafood	2	0.01
climbing	2	0.01	cauliflower	2	0.01
runner	2	0.01	clues	2	0.01
Japan	2	0.01	deep	2	0.01
high-quality	2	0.01	increase	2	0.01
European	2	0.01	corn	2	0.01
safer	2	0.01	cup	2	0.01

thickly	2	0.01	competitions	2	0.01
medium	2	0.01	joining	2	0.01
rounds	2	0.01	prizes	2	0.01
vanilla	2	0.01	registration	2	0.01
custard	2	0.01	fee	2	0.01
sides	2	0.01	competition	2	0.01
stick	2	0.01	at it	2	0.01
British	2	0.01	photographs	2	0.01
chocolate	2	0.01	hotel	2	0.01
steak	2	0.01	Sumela	2	0.01
nothing	2	0.01	Pisa	2	0.01
square	2	0.01	cultures	2	0.01
cake	2	0.01	cultural	2	0.01
perfect	2	0.01	sembusek	2	0.01
sign language	2	0.01	harire	2	0.01
German	2	0.01	American	2	0.01
symbol	2	0.01	weed	2	0.01
speakers	2	0.01	Belinda	2	0.01
calls on	2	0.01	timetable	2	0.01
landline	2	0.01	Frankie	2	0.01
e.	2	0.01	finish	2	0.01
g.	2	0.01	arrive	2	0.01
storage	2	0.01	lunch break	2	0.01
communicator	2	0.01	task	2	0.01
rh-10	2	0.01	materials	2	0.01
product	2	0.01	rule	2	0.01
human	2	0.01	physicist	2	0.01
flash	2	0.01	cosmonaut	2	0.01
on the other side	2	0.01	engineer	2	0.01
uncomfortable	2	0.01	computer programmer	2	0.01
identities	2	0.01	cell	2	0.01
known	2	0.01	knew	2	0.01
regions	2	0.01	experiment	2	0.01
middle east	2	0.01	Tyler Hadzicki	2	0.01
total	2	0.01	market	2	0.01
south America	2	0.01	presented	2	0.01
habit	2	0.01	finalized	2	0.01
items	2	0.01	closer	2	0.01
heart-stopping	2	0.01	Tyler	2	0.01
basic	2	0.01	impact	2	0.01
white-water rafting	2	0.01	effect	2	0.01
feet	2	0.01	fire	2	0.01

storm	2	0.01		parties	1	0.00
eruption	2	0.01		concerts	1	0.00
slide	2	0.01		ad	1	0.00
agency	2	0.01		Frank Frank	1	0.00
Everest	2	0.01		Suzy Kevin Ahmet	1	0.00
climbers	2	0.01		Friday night	1	0.00
chili	2	0.01		re	1	0.00
eco-friendly	2	0.01		fine	1	0.00
short recorded	2	0.01		hair	1	0.00
everyday	2	0.01		arrangements	1	0.00
offer	2	0.01		Eskiehir	1	0.00
apologize	2	0.01		Salim	1	0.00
dislike	2	0.01		Fatih	1	0.00
willingness	2	0.01		partner example	1	0.00
advertisements	2	0.01		rock concert	1	0.00
written	2	0.01		computer game	1	0.00
travel	2	0.01		Jane Dave Nancy	1	0.00
participle	2	0.01		role	1	0.00
bought	2	0.01		roles	1	0.00
felt	2	0.01		make sure	1	0.00
found	2	0.01		beginning	1	0.00
got	2	0.01		invites out	1	0.00
held	2	0.01		excuse example	1	0.00
kept	2	0.01		November	1	0.00
led	2	0.01		155	1	0.00
lost	2	0.01		Carmen	1	0.00
meant	2	0.01		Ave	1	0.00
met	2	0.01		Birmingham	1	0.00
paid	2	0.01		get off	1	0.00
sent	2	0.01		Forest Road	1	0.00
sat	2	0.01		cheers sally	1	0.00
spent	2	0.01		par-tea	1	0.00
stood	2	0.01		appears	1	0.00
told	2	0.01		request	1	0.00
thought	2	0.01		proverb	1	0.00
understood	2	0.01		mind	1	0.00
adress	2	0.01		in two	1	0.00
extraordinary	2	0.01		bodies	1	0.00
Burj Dubai	2	0.00		depend	1	0.00
ikbebet	2	0.00		interests	1	0.00
parties with	1	0.00		don	1	0.00
going for	1	0.00		closest	1	0.00

argue	1	0.00	oh dear	1	0.00
most of the time	1	0.00	class work	1	0.00
friendships	1	0.00	hearth	1	0.00
ourselves	1	0.00	RIMMS	1	0.00
dialogs	1	0.00	Oliver CARTY	1	0.00
Donna Gale	1	0.00	Kevin ALLBRIGHT	1	0.00
come in	1	0.00	players	1	0.00
Gale Gale	1	0.00	titles	1	0.00
cookies	1	0.00	methods	1	0.00
hungry	1	0.00	potatoes	1	0.00
collection	1	0.00	tasty	1	0.00
in your mind	1	0.00	next three	1	0.00
frozen	1	0.00	cooking look	1	0.00
they say	1	0.00	package	1	0.00
choice	1	0.00	lightly	1	0.00
slave	1	0.00	tea towel	1	0.00
watch it	1	0.00	pilaf work	1	0.00
apologies	1	0.00	healthy	1	0.00
welcoming	1	0.00	belong	1	0.00
coming in	1	0.00	kabuli	1	0.00
chatting	1	0.00	palaw	1	0.00
seeing	1	0.00	steamed	1	0.00
ridiculous	1	0.00	Afghanistan	1	0.00
bike.	1	0.00	Turkmenistan	1	0.00
father work	1	0.00	fingers	1	0.00
middle school	1	0.00	Turkmen	1	0.00
takes	1	0.00	knives	1	0.00
midnight	1	0.00	onion sauce	1	0.00
adverbs	1	0.00	round	1	0.00
Saturdays	1	0.00	lamb soup	1	0.00
meets	1	0.00	Amin	1	0.00
stay up	1	0.00	God	1	0.00
life look	1	0.00	leeks	1	0.00
gaps	1	0.00	boiled	1	0.00
rap	1	0.00	Kabuli Palaw Beshbarmak	1	0.00
electronic	1	0.00	take a look	1	0.00
heavy metal	1	0.00	probable	1	0.00
hip	1	0.00	pre-heat	1	0.00
hop	1	0.00	wet	1	0.00
pen	1	0.00	cups	1	0.00
three times a week	1	0.00	Berry Muffins	1	0.00
untidy	1	0.00	Mr. Rick	1	0.00

Nelly	1	0.00	screen work	1	0.00
informal	1	0.00	software	1	0.00
voicemails	1	0.00	asks	1	0.00
Amanda Betty Victor	1	0.00	reset	1	0.00
partner work	1	0.00	settings	1	0.00
calls	1	0.00	device	1	0.00
wants	1	0.00	affects	1	0.00
John speaking	1	0.00	hobby	1	0.00
call back	1	0.00	everyone	1	0.00
Rosa	1	0.00	loves	1	0.00
suppose	1	0.00	experts	1	0.00
Eric express	1	0.00	pcs	1	0.00
chess	1	0.00	laptops	1	0.00
used	1	0.00	tablets	1	0.00
SMS Short	1	0.00	smartphones	1	0.00
Britain	1	0.00	surely	1	0.00
combinations	1	0.00	classmate	1	0.00
symbols	1	0.00	Jen	1	0.00
worried	1	0.00	Tim Ken	1	0.00
negative	1	0.00	z	1	0.00
cul8r	1	0.00	anyway	1	0.00
2nite	1	0.00	Ken	1	0.00
lol	1	0.00	results	1	0.00
BBL	1	0.00	Hasan	1	0.00
be back	1	0.00	checks	1	0.00
communication work	1	0.00	once a day	1	0.00
dial	1	0.00	mini golf	1	0.00
phone bill	1	0.00	Helen who	1	0.00
Bill speaking	1	0.00	taking part	1	0.00
Marc Bill	1	0.00	activites	1	0.00
bowling	1	0.00	involving	1	0.00
bowling alley	1	0.00	physical	1	0.00
Palm Street Marc	1	0.00	Helen Snow	1	0.00
services	1	0.00	all over the world	1	0.00
username	1	0.00	attracts	1	0.00
computer file	1	0.00	national water	1	0.00
videos	1	0.00	centre	1	0.00
Jason Seville	1	0.00	Joe Pixy	1	0.00
connecting	1	0.00	Helen Joe Joe	1	0.00
reads	1	0.00	indoor	1	0.00
newspapers	1	0.00	course	1	0.00
laptop	1	0.00	outdoors	1	0.00

challenge	1	0.00		St. Paul	1	0.00
push	1	0.00		prophet	1	0.00
vacation work	1	0.00		Daniel	1	0.00
d	1	0.00		eshab-ı Kehf	1	0.00
summer holidays	1	0.00		Cave Roman	1	0.00
driving	1	0.00		Road	1	0.00
healthier	1	0.00		k1rkka1k	1	0.00
jumps	1	0.00		Bedesten Tarsus	1	0.00
aim	1	0.00		recreation	1	0.00
runners	1	0.00		safari	1	0.00
high-quality	1	0.00		off-road	1	0.00
weigh	1	0.00		kebabs	1	0.00
faster	1	0.00		Roman Road	1	0.00
Nevşehir	1	0.00		Cleopatra	1	0.00
mula	1	0.00		islands	1	0.00
full	1	0.00		travel brochure	1	0.00
Paris Paris	1	0.00		come from	1	0.00
Moscow Russia Paris	1	0.00		take turns	1	0.00
incredible	1	0.00		Kingston	1	0.00
lovely	1	0.00		English total	1	0.00
tower bridge	1	0.00		Jamaican	1	0.00
Kremlin palace	1	0.00		smiling	1	0.00
Red Square	1	0.00		crystal-clear	1	0.00
pages	1	0.00		Caribbean Jamaica	1	0.00
convenient	1	0.00		Caribbean islands	1	0.00
south	1	0.00		English-speaking	1	0.00
20	1	0.00		Caribbean Sea	1	0.00
inland	1	0.00		comes from	1	0.00
going back	1	0.00		Xaymaca	1	0.00
traders	1	0.00		wood	1	0.00
focal	1	0.00		springs	1	0.00
empire	1	0.00		along	1	0.00
Cilicia	1	0.00		Marmara coast	1	0.00
meeting	1	0.00		North western Anatolia	1	0.00
Mark Antony	1	0.00		Emir	1	0.00
Cleopatra	1	0.00		Sultan mosque	1	0.00
junction	1	0.00		If you like	1	0.00
Cilician	1	0.00		Uludağ	1	0.00
Çukurova	1	0.00		Turkey Bursa	1	0.00
Anatolia	1	0.00		rich	1	0.00
Cleopatra	1	0.00		iskender	1	0.00
Justinian_	1	0.00		inegöl	1	0.00

meatball	1	0.00	taking out	1	0.00
Kemalpaşa	1	0.00	Diane Diane Curtis	1	0.00
candied	1	0.00	in my opinion	1	0.00
chestnut	1	0.00	titled	1	0.00
easily	1	0.00	explains	1	0.00
hotels	1	0.00	writes	1	0.00
hostels	1	0.00	Jim Watson	1	0.00
extra	1	0.00	Stephanie Kwolek	1	0.00
family work	1	0.00	James Watson	1	0.00
cat	1	0.00	Stephanie Kwolek	1	0.00
sharing	1	0.00	James Watson Stephanie	1	0.00
Sunday today	1	0.00	Charles Babbage	1	0.00
sweeping	1	0.00	Benjamin Franklin	1	0.00
leaves	1	0.00	Galileo	1	0.00
chairs	1	0.00	invent/discover	1	0.00
knows	1	0.00	humanity work	1	0.00
living room	1	0.00	assistant	1	0.00
hard work	1	0.00	Professor	1	0.00
duty	1	0.00	Utkan Demirci	1	0.00
enjoying	1	0.00	a number of	1	0.00
waiting	1	0.00	micro	1	0.00
vacuuming	1	0.00	diagnoses	1	0.00
each day	1	0.00	HIV	1	0.00
bathroom floor	1	0.00	patients	1	0.00
schoolwork	1	0.00	Tanzania	1	0.00
faces	1	0.00	awards	1	0.00
listener	1	0.00	medial	1	0.00
take-off	1	0.00	Harvard	1	0.00
do your best	1	0.00	medical school	1	0.00
leaving	1	0.00	young investigator	1	0.00
neat	1	0.00	these days	1	0.00
jump on	1	0.00	cure	1	0.00
couches	1	0.00	common cold	1	0.00
quiet	1	0.00	organisms	1	0.00
siblings	1	0.00	milligram	1	0.00
rules classroom	1	0.00	printed	1	0.00
guesses	1	0.00	nowadays	1	0.00
Diane Curtis Nina	1	0.00	MAYA	1	0.00
his/her	1	0.00	perfected	1	0.00
bird	1	0.00	amounts	1	0.00
watering	1	0.00	standard Nanomedicine	1	0.00
takes care of	1	0.00	usage	1	0.00

endless	1	0.00	beauty	1	0.00
drugs	1	0.00	oceans	1	0.00
body	1	0.00	covering	1	0.00
microscopic	1	0.00	itself	1	0.00
robots	1	0.00	upside	1	0.00
repair	1	0.00	spinning	1	0.00
damaged	1	0.00	able	1	0.00
tissues	1	0.00	across	1	0.00
hunt	1	0.00	falling	1	0.00
certain	1	0.00	side	1	0.00
cells	1	0.00	causing	1	0.00
currently	1	0.00	moving	1	0.00
in its infancy	1	0.00	suddenly	1	0.00
since	1	0.00	steep	1	0.00
conducting	1	0.00	slope	1	0.00
doctor	1	0.00	took place	1	0.00
vaccinating	1	0.00	northwest	1	0.00
searching	1	0.00	Indonesian	1	0.00
heating	1	0.00	Sumatra	1	0.00
researchers	1	0.00	formed	1	0.00
fossils	1	0.00	travelled	1	0.00
excavation	1	0.00	speeds	1	0.00
dentist	1	0.00	as far as	1	0.00
x-ray	1	0.00	away	1	0.00
operating	1	0.00	affected	1	0.00
machine	1	0.00	eleven	1	0.00
particularly	1	0.00	hardest	1	0.00
programs	1	0.00	Indonesia Sri Lanka	1	0.00
missed the point	1	0.00	Maldives more	1	0.00
accelerating	1	0.00	thousands	1	0.00
Perito Moreno Glacier	1	0.00	millions	1	0.00
Los Glaciares	1	0.00	destructive	1	0.00
Santa Cruz	1	0.00	tsunami	1	0.00
Museums	1	0.00	harbor	1	0.00
exhibitions	1	0.00	tsunamis	1	0.00
principles	1	0.00	eruptions	1	0.00
stage	1	0.00	landslides	1	0.00
hidden	1	0.00	crashing	1	0.00
film	1	0.00	news work	1	0.00
changed	1	0.00	stroke	1	0.00
stunning	1	0.00	city of Mexico	1	0.00
creatures	1	0.00	Kabul Afghanistan	1	0.00

moved	1	0.00	rainfall	1	0.00
kilometres	1	0.00	turn	1	0.00
hits	1	0.00	too fast	1	0.00
face	1	0.00	dams	1	0.00
bookcases	1	0.00	scripts	1	0.00
button	1	0.00	audio	1	0.00
get-out	1	0.00	Mike Frank	1	0.00
systems	1	0.00	Abby Suzy	1	0.00
fire alarms	1	0.00	start Abby	1	0.00
activate	1	0.00	lasts	1	0.00
mouth	1	0.00	Suzy will	1	0.00
shirt	1	0.00	pick me up	1	0.00
pipe	1	0.00	Kevin Kevin	1	0.00
object	1	0.00	says	1	0.00
unless	1	0.00	nothing special	1	0.00
inhaling	1	0.00	seems	1	0.00
quantities	1	0.00	Ashley Jane Ashley	1	0.00
aftershocks	1	0.00	Ashley Jane Jane	1	0.00
bring down	1	0.00	Pete Dave Pete	1	0.00
wires	1	0.00	Pete Dave	1	0.00
on hold	1	0.00	Say Pete	1	0.00
on to	1	0.00	studies	1	0.00
firmly	1	0.00	Helen Nancy Helen	1	0.00
be prepared to	1	0.00	Helen Nancy Nancy	1	0.00
interior	1	0.00	nature club	1	0.00
neck	1	0.00	Helen on Sunday	1	0.00
earth work	1	0.00	10	1	0.00
else	1	0.00	Guests	1	0.00
window	1	0.00	matter	1	0.00
eco	1	0.00	chain gear	1	0.00
labelled	1	0.00	Kensington middle school	1	0.00
detergents	1	0.00	camping	1	0.00
eco conscious	1	0.00	jogging	1	0.00
radar	1	0.00	gears	1	0.00
top priority	1	0.00	comfortable	1	0.00
launch	1	0.00	program	1	0.00
campaigns	1	0.00	legs	1	0.00
occurs	1	0.00	stronger	1	0.00
period of time	1	0.00	brakes	1	0.00
fires	1	0.00	gear	1	0.00
animals	1	0.00	pedals	1	0.00
months	1	0.00	tyres	1	0.00

Kate Kate	1	0.00		pair	1	0.00
liter	1	0.00		trainers	1	0.00
pot	1	0.00		stairs	1	0.00
packet	1	0.00		trees	1	0.00
few times	1	0.00		scared	1	0.00
for a minute	1	0.00		years ago	1	0.00
plate	1	0.00		increasing	1	0.00
gently	1	0.00		popularity	1	0.00
seconds	1	0.00		look like	1	0.00
popping	1	0.00		toys	1	0.00
crack	1	0.00		one fourth	1	0.00
beat	1	0.00		normal	1	0.00
completely	1	0.00		weighs	1	0.00
diced	1	0.00		kilos	1	0.00
bon	1	0.00		reach	1	0.00
appetite	1	0.00		slower	1	0.00
vwill	1	0.00		such a	1	0.00
Betty	1	0.00		vehicle	1	0.00
Victor	1	0.00		motorcyclists	1	0.00
grandmother	1	0.00		careers	1	0.00
eightieth	1	0.00		grand prix	1	0.00
Rosa speaking	1	0.00		champion	1	0.00
Internet Jason Jason	1	0.00		Valentine Rossi	1	0.00
jasonseville	1	0.00		rode	1	0.00
l	1	0.00		eight	1	0.00
www	1	0.00		Agra India	1	0.00
Danny who	1	0.00		Shah Jahan	1	0.00
accesses	1	0.00		wife	1	0.00
cafs	1	0.00		Mumtaz Mahal	1	0.00
street	1	0.00		Persian	1	0.00
countryside	1	0.00		Indian	1	0.00
open spaces	1	0.00		waterfalls	1	0.00
becoming	1	0.00		Canadian	1	0.00
France	1	0.00		falls	1	0.00
as fast as possible	1	0.00		sound	1	0.00
whatever	1	0.00		goes over	1	0.00
in the way	1	0.00		lands	1	0.00
parkour class	1	0.00		bottom	1	0.00
Jack Sanders Jack	1	0.00		miracle	1	0.00
experienced	1	0.00		Ahmet the first	1	0.00
fit	1	0.00		Mehmet Ağa	1	0.00
equipment	1	0.00		clock towers	1	0.00

Ben in London	1	0.00	bringing	1	0.00
holds	1	0.00	frightened	1	0.00
icon	1	0.00	streets	1	0.00
pharaoh	1	0.00	depth	1	0.00
Khufu	1	0.00	according	1	0.00
constructed	1	0.00	Kabul Afghanistan	1	0.00
Sally Sally	1	0.00	United nations	1	0.00
ever	1	0.00	On the ground	1	0.00
wow	1	0.00	reached	1	0.00
such	1	0.00	hospitals	1	0.00
topkapı	1	0.00	treated	1	0.00
Bosphorus	1	0.00	officials	1	0.00
those	1	0.00	elementary school	1	0.00
went with	1	0.00	worst-hit	1	0.00
cousins	1	0.00	workbook	1	0.00
sunny	1	0.00	contents	1	0.00
blog	1	0.00	showing	1	0.00
Mother Children	1	0.00	CnM	1	0.00
Mother Leo	1	0.00	theaters	1	0.00
cupboard	1	0.00	Sunday evening	1	0.00
Mother Betsy	1	0.00	at all	1	0.00
Deal Denis	1	0.00	tigers	1	0.00
building blocks	1	0.00	game show	1	0.00
Diane	1	0.00	shall	1	0.00
Nina	1	0.00	sounds like	1	0.00
birds	1	0.00	Samuel	1	0.00
seem	1	0.00	surprise	1	0.00
brief	1	0.00	Henry Kate	1	0.00
helped	1	0.00	already	1	0.00
range	1	0.00	planned	1	0.00
lifting up	1	0.00	cds	1	0.00
mines	1	0.00	archive	1	0.00
wells	1	0.00	got it	1	0.00
Alexander Bell	1	0.00	sending	1	0.00
Benjamin Franklin	1	0.00	card	1	0.00
Galileo	1	0.00	invited	1	0.00
revolutionary	1	0.00	action movie	1	0.00
theories	1	0.00	Ellen	1	0.00
good afternoon	1	0.00	action	1	0.00
viewers	1	0.00	concert	1	0.00
Angela star	1	0.00	graduation	1	0.00
Frank Donna	1	0.00	feel like	1	0.00

Glen	1	0.00		Oakland high school	1	0.00
promise	1	0.00		Vancouver	1	0.00
depressed	1	0.00		7_o'clock	1	0.00
feel free	1	0.00		subject	1	0.00
gonna	1	0.00		geography	1	0.00
bear	1	0.00		badminton	1	0.00
tiny	1	0.00		drums	1	0.00
bug	1	0.00		crackers	1	0.00
I swear to God	1	0.00		hip-hop	1	0.00
Hugs	1	0.00		dramas	1	0.00
lonely	1	0.00		musicals	1	0.00
Nicole George	1	0.00		deals	1	0.00
glad	1	0.00		wishes	1	0.00
restaurant	1	0.00		Rob Spencer	1	0.00
hairdresser	1	0.00		style	1	0.00
barber	1	0.00		zucchinis	1	0.00
half past seven	1	0.00		peppers	1	0.00
Ben	1	0.00		cabbages	1	0.00
tea	1	0.00		garlic	1	0.00
choir	1	0.00		softened	1	0.00
art class	1	0.00		3	1	0.00
rooms	1	0.00		tender	1	0.00
very much	1	0.00		Blend	1	0.00
sentence	1	0.00		jug	1	0.00
category	1	0.00		stock	1	0.00
orchestra	1	0.00		sauce	1	0.00
to be honest	1	0.00		boils	1	0.00
game	1	0.00		minute	1	0.00
practicing	1	0.00		season	1	0.00
martial arts	1	0.00		cloves	1	0.00
tutor	1	0.00		crushed	1	0.00
Megan	1	0.00		seeded	1	0.00
cannot stand	1	0.00		florets	1	0.00
Working out	1	0.00		lengths	1	0.00
teach	1	0.00		peeled	1	0.00
next month	1	0.00		melted	1	0.00
keep fit	1	0.00		ml	1	0.00
coach	1	0.00		strawberry	1	0.00
best wishes	1	0.00		warmed	1	0.00
letter go	1	0.00		beaten	1	0.00
Oakland Road	1	0.00		preheat	1	0.00
Sezen	1	0.00		mix together	1	0.00

soft	1	0.00	Chinese	1	0.00
sticky	1	0.00	originally	1	0.00
rectangle	1	0.00	rewrite	1	0.00
all over	1	0.00	texting	1	0.00
sprinkle	1	0.00	dominates	1	0.00
greased	1	0.00	general	1	0.00
baking sheet	1	0.00	just about	1	0.00
fasten	1	0.00	dominant	1	0.00
mentioned	1	0.00	mode	1	0.00
fish and chip shops	1	0.00	socialize	1	0.00
roast beef	1	0.00	outside of	1	0.00
pudding	1	0.00	on a daily basis	1	0.00
cattle	1	0.00	landlines with	1	0.00
sheep	1	0.00	interest	1	0.00
salad	1	0.00	exchanging	1	0.00
chips	1	0.00	text messaging	1	0.00
prices	1	0.00	gadget	1	0.00
fish and chips	1	0.00	technical	1	0.00
meat pie	1	0.00	specifications	1	0.00
however	1	0.00	translate	1	0.00
throughout	1	0.00	pronounce	1	0.00
whole	1	0.00	digital	1	0.00
barbeques	1	0.00	translator	1	0.00
parks	1	0.00	fits	1	0.00
backyards	1	0.00	packs	1	0.00
sweet tooth	1	0.00	inch	1	0.00
Lamington	1	0.00	wqvga	1	0.00
covered	1	0.00	qwerty	1	0.00
Anzac	1	0.00	GB	1	0.00
biscuit	1	0.00	translates	1	0.00
oats	1	0.00	contains	1	0.00
Pavlova	1	0.00	displays	1	0.00
Merhaba	1	0.00	recognizes	1	0.00
Bonjour	1	0.00	handwriting	1	0.00
Guten tag	1	0.00	touchscreen	1	0.00
ol	1	0.00	pronounces	1	0.00
Emma	1	0.00	parts with	1	0.00
code	1	0.00	usb	1	0.00
alphabet	1	0.00	DVD	1	0.00
Stacy	1	0.00	discs	1	0.00
person speaking	1	0.00	printer	1	0.00
Russian	1	0.00	microphone	1	0.00

mouse	1	0.00	in other words	1	0.00
webcam	1	0.00	stands for	1	0.00
puzzle	1	0.00	antenna	1	0.00
send in	1	0.00	span	1	0.00
spends	1	0.00	jumping off	1	0.00
surfing	1	0.00	towers	1	0.00
put on	1	0.00	bridges	1	0.00
avoid	1	0.00	of course	1	0.00
anyone	1	0.00	basics	1	0.00
except	1	0.00	difference	1	0.00
signs in	1	0.00	death	1	0.00
identity	1	0.00	ankle	1	0.00
careful	1	0.00	protection	1	0.00
strangers	1	0.00	touch	1	0.00
in-person	1	0.00	proper	1	0.00
age	1	0.00	footwear	1	0.00
adults	1	0.00	ankles	1	0.00
rude	1	0.00	medical	1	0.00
you see	1	0.00	goes wrong	1	0.00
behavior	1	0.00	skis	1	0.00
source	1	0.00	snowboard	1	0.00
geographical	1	0.00	Adeline Sparks	1	0.00
north	1	0.00	festival	1	0.00
distributions	1	0.00	sparksadeline	1	0.00
Asia	1	0.00	mail	1	0.00
Europe	1	0.00	Missouri	1	0.00
North America	1	0.00	t-shirt	1	0.00
million million	1	0.00	raffle	1	0.00
projects	1	0.00	coupon	1	0.00
serials	1	0.00	in-line	1	0.00
look for	1	0.00	skating	1	0.00
adventurous	1	0.00	go-karting	1	0.00
least	1	0.00	souvenirs	1	0.00
trekking	1	0.00	economy class	1	0.00
rock climbing	1	0.00	postcards	1	0.00
6	1	0.00	wonderful	1	0.00
Medical treatment	1	0.00	stayed	1	0.00
hat	1	0.00	weeks	1	0.00
harmful	1	0.00	joined	1	0.00
shocking	1	0.00	boat	1	0.00
hurt	1	0.00	tours	1	0.00
pain	1	0.00	7	1	0.00

skyscraper	1	0.00		Seljuk	1	0.00
United Arab Emirates	1	0.00		Turks	1	0.00
Himeji Castle	1	0.00		mosaics	1	0.00
finest	1	0.00		crossroad	1	0.00
Zigana	1	0.00		religions	1	0.00
Trabzon Turkey	1	0.00		mosques	1	0.00
Meryem Ana	1	0.00		churches	1	0.00
Virgin Mary	1	0.00		madrassas	1	0.00
Great wall	1	0.00		Virgin Mary Church	1	0.00
greatest	1	0.00		well known	1	0.00
stretches	1	0.00		recent	1	0.00
east	1	0.00		attraction	1	0.00
west of China	1	0.00		candidate	1	0.00
bell tower	1	0.00		UNESCO	1	0.00
lean	1	0.00		Heritage	1	0.00
leans	1	0.00		Say Gale Sanders	1	0.00
degrees	1	0.00		USA Gale	1	0.00
London Eye	1	0.00		charts	1	0.00
Ferris	1	0.00		Bill Bill	1	0.00
south bank	1	0.00		father bill	1	0.00
River Thames in London	1	0.00		mother bill	1	0.00
entire	1	0.00		Linda Linda	1	0.00
diameter	1	0.00		live with	1	0.00
discovering	1	0.00		our father	1	0.00
learning	1	0.00		loads	1	0.00
south-eastern	1	0.00		vacuums	1	0.00
Anatolian	1	0.00		every Saturday	1	0.00
continental	1	0.00		flat	1	0.00
summer months	1	0.00		apartment block	1	0.00
temperatures	1	0.00		every Wednesday afternoon	1	0.00
oldest	1	0.00		emptying	1	0.00
settled	1	0.00		canteen	1	0.00
areas	1	0.00		packed lunch	1	0.00
upper	1	0.00		Thursday on Wednesday	1	0.00
Mesopotamia	1	0.00		music club	1	0.00
Babylonians	1	0.00		rock band	1	0.00
Hittites	1	0.00		rocky	1	0.00
Assyrians	1	0.00		mistakes	1	0.00
romans	1	0.00		focus	1	0.00
Byzantines	1	0.00		cooperate	1	0.00
Abbasid	1	0.00		group	1	0.00
Caliphate	1	0.00		disorganized	1	0.00

rubbish in	1	0.00	viewing	1	0.00
Charles Babbage Thomas	1	0.00	specimens	1	0.00
women	1	0.00	microscopes	1	0.00
geneticist	1	0.00	looking into	1	0.00
polish	1	0.00	craters	1	0.00
radioactivity	1	0.00	moon	1	0.00
polonium	1	0.00	underline	1	0.00
radium	1	0.00	sits	1	0.00
first-person	1	0.00	bigger	1	0.00
Nobel	1	0.00	riders	1	0.00
physics	1	0.00	reinvent	1	0.00
chemistry	1	0.00	year-old	1	0.00
soviet	1	0.00	named	1	0.00
woman	1	0.00	somewhat	1	0.00
civilian	1	0.00	front	1	0.00
carried	1	0.00	rear	1	0.00
olympic	1	0.00	smaller	1	0.00
flag	1	0.00	lock	1	0.00
opening	1	0.00	Tyler first	1	0.00
ceremony	1	0.00	sponsors	1	0.00
Olympics	1	0.00	raised	1	0.00
geneticists	1	0.00	makes it	1	0.00
worked	1	0.00	biker	1	0.00
function	1	0.00	heights	1	0.00
chromosomes	1	0.00	thousand	1	0.00
received	1	0.00	on the road	1	0.00
writer	1	0.00	as for	1	0.00
algorithm	1	0.00	reaches	1	0.00
Ada Lovelace Marie	1	0.00	sponsor	1	0.00
test tubes	1	0.00	advanced	1	0.00
observe	1	0.00	likely	1	0.00
reactions	1	0.00	organization	1	0.00
alternative	1	0.00	money	1	0.00
sources	1	0.00	making it	1	0.00
model	1	0.00	unscramble	1	0.00
turbine	1	0.00	tornado hurricane	1	0.00
geology	1	0.00	uncontrolled	1	0.00
soil	1	0.00	occurring	1	0.00
samples	1	0.00	lightning	1	0.00
testing	1	0.00	explosion	1	0.00
efficiency	1	0.00	volcano	1	0.00
lab	1	0.00	blizzard	1	0.00

winds	1	0.00		bottles	1	0.00
men	1	0.00		limit	1	0.00
Colorado	1	0.00		disposable	1	0.00
witness	1	0.00		tree	1	0.00
sounded like	1	0.00		both sides	1	0.00
Taiwan	1	0.00		instead of	1	0.00
emergency	1	0.00		conserve	1	0.00
management	1	0.00		appliances	1	0.00
rush	1	0.00		in use	1	0.00
out of	1	0.00		lights	1	0.00
struck on	1	0.00		portfolio	1	0.00
base camp	1	0.00		surname	1	0.00
rescued	1	0.00		grade class	1	0.00
southern	1	0.00		page	1	0.00
xinhua	1	0.00		familiar	1	0.00
residents	1	0.00		exhibit	1	0.00
had to	1	0.00		apply	1	0.00
traffic	1	0.00		skills	1	0.00
telecommunication	1	0.00		situations	1	0.00
map	1	0.00		similarities	1	0.00
TURKEY CHINA JAPAN	1	0.00		differences	1	0.00
leading	1	0.00		movies free	1	0.00
cost billion	1	0.00		home life	1	0.00
deaths	1	0.00		processed	1	0.00
worth	1	0.00		presentation	1	0.00
environment	1	0.00		prepared	1	0.00
switch off	1	0.00		linked	1	0.00
none	1	0.00		context	1	0.00
household	1	0.00		diversity	1	0.00
as much as	1	0.00		readings	1	0.00
buying	1	0.00		responding	1	0.00
matters	1	0.00		containing	1	0.00
cost	1	0.00		seek	1	0.00
many times	1	0.00		opportunities	1	0.00
bag	1	0.00		predictable	1	0.00
drives	1	0.00		requests	1	0.00
picture	1	0.00		express myself	1	0.00
recycling	1	0.00		several	1	0.00
bins	1	0.00		intended	1	0.00
sort	1	0.00		meaning	1	0.00
aluminum	1	0.00		expression	1	0.00
cans	1	0.00		provided	1	0.00

descriptive	1	0.00		attachment book	1	0.00
features	1	0.00		modem classmate	1	0.00
particular	1	0.00		in touch	1	0.00
maps	1	0.00		comparison	1	0.00
advertisement	1	0.00		fan	1	0.00
maintain	1	0.00		manufacturer	1	0.00
concentration	1	0.00		narrative	1	0.00
motivation	1	0.00		architectural	1	0.00
changes	1	0.00		cave	1	0.00
content	1	0.00		civilization	1	0.00
diaries	1	0.00		construct	1	0.00
journal	1	0.00		empire	1	0.00
poems	1	0.00		hostel	1	0.00
report	1	0.00		trader	1	0.00
key	1	0.00		waterfall	1	0.00
in short	1	0.00		worker	1	0.00
phrase	1	0.00		couch	1	0.00
endeavor	1	0.00		obligation	1	0.00
understanding	1	0.00		sibling	1	0.00
summary	1	0.00		pet	1	0.00
buddy	1	0.00		accelerate	1	0.00
apology	1	0.00		cancer	1	0.00
organize	1	0.00		conduct	1	0.00
pick	1	0.00		drug	1	0.00
response	1	0.00		ecosystem	1	0.00
asleep	1	0.00		worming	1	0.00
fun gear	1	0.00		humanity	1	0.00
musical instruments	1	0.00		investigator	1	0.00
pancake	1	0.00		light bulb	1	0.00
ticket	1	0.00		micro-device	1	0.00
absorb	1	0.00		operate	1	0.00
berry	1	0.00		organism	1	0.00
egg	1	0.00		patient	1	0.00
eggplant	1	0.00		principle	1	0.00
method	1	0.00		surgery	1	0.00
mushroom	1	0.00		vaccinate	1	0.00
potato	1	0.00		aftershock	1	0.00
salt saucepan	1	0.00		authority	1	0.00
abbreviation	1	0.00		burn	1	0.00
hang on	1	0.00		campaign	1	0.00
mobile phone	1	0.00		crash	1	0.00
worry	1	0.00		dam	1	0.00

double-sided	1	0.00	saw	1	0.00
eco-conscious	1	0.00	seen	1	0.00
wire	1	0.00	shown	1	0.00
frighten	1	0.00	spoke	1	0.00
inhale	1	0.00	stand	1	0.00
injury	1	0.00	wore	1	0.00
suggestion	1	0.00	worn	1	0.00
priority	1	0.00	wrote	1	0.00
safe-room	1	0.00	tea party	1	0.00
spin	1	0.00	phone number	1	0.00
system	1	0.00	e-mail	1	0.00
stay away	1	0.00	Liz	1	0.00
strike	1	0.00	Sonya Hillary	1	0.00
suburb	1	0.00	Darren	1	0.00
upside-down	1	0.00	tennis club	1	0.00
irregular	1	0.00	Friday afternoon	1	0.00
begun	1	0.00	rackets	1	0.00
broke	1	0.00	cheers Mary	1	0.00
chose	1	0.00	Eric Jeff	1	0.00
chosen	1	0.00	let know	1	0.00
came	1	0.00	journey	1	0.00
drew	1	0.00	Khan Steve Carell	1	0.00
drawn	1	0.00	schmucks	1	0.00
drove	1	0.00	director	1	0.00
driven	1	0.00	outrageously	1	0.00
fell	1	0.00	funny	1	0.00
fallen	1	0.00	let go	1	0.00
gotten	1	0.00	gravity	1	0.00
gone	1	0.00	Tom Cruise	1	0.00
grow	1	0.00	Jack Reacher	1	0.00
grew	1	0.00	gift	1	0.00
grown	1	0.00	Scott Harfield Little	1	0.00
lead	1	0.00	Searles	1	0.00
lay	1	0.00	Basildon Essex	1	0.00
lain	1	0.00	coming over	1	0.00
lose	1	0.00	guys	1	0.00
pay	1	0.00	wish	1	0.00
rise	1	0.00	hereexcept	1	0.00
rose	1	0.00	pyramids	1	0.00
risen	1	0.00	Alice	1	0.00
ran	1	0.00			

APPENDIX E

THE COMPARISON OF THE FREQUENCY OF THE WORDS IN THE

COURSEBOOK WITH THE BRITISH NATIONAL CORPUS

* the words which have the Log-Likelihood value more than 50.00 have been given

Item	O1	%1	O2	%2		LL	LogRatio
unit	194	0.69	38	0.00	+	1187.02	7.49
internet	99	0.35	0	0.00	+	710.24	-7.14
write	136	0.49	226	0.02	+	509.15	4.40
sentences	63	0.23	2	0.00	+	434.22	10.11
your	348	1.24	3250	0.33	+	392.05	1.91
activity	68	0.24	29	0.00	+	371.13	6.36
text	63	0.23	20	0.00	+	361.43	6.79
the	1749	6.25	37283	3.79	+	359.71	0.72
photos	52	0.19	2	0.00	+	356.06	9.83
below	69	0.25	42	0.00	+	350.13	5.85
lesson	60	0.21	30	0.00	+	317.56	6.13
read	94	0.34	229	0.02	+	297.65	3.85
science	52	0.19	19	0.00	+	291.64	6.59
questions	76	0.27	144	0.01	+	269.70	4.21
am	94	0.34	281	0.03	+	267.85	3.55
classmates	36	0.13	0	0.00	+	258.27	-8.60
check	66	0.24	106	0.01	+	250.39	4.45
chores	37	0.13	3	0.00	+	244.30	8.76
simple	55	0.20	59	0.01	+	239.99	5.03
example	52	0.19	49	0.00	+	235.88	5.22
favorite	32	0.11	0	0.00	+	229.57	-8.77
movies	33	0.12	1	0.00	+	227.78	10.18
share	57	0.20	87	0.01	+	220.48	4.52
sports	41	0.15	20	0.00	+	218.08	6.17
online	35	0.13	8	0.00	+	210.23	7.26
usually	65	0.23	171	0.02	+	198.11	3.74
partner	37	0.13	18	0.00	+	196.91	6.17
use	107	0.38	614	0.06	+	196.56	2.61
my	215	0.77	2354	0.24	+	196.44	1.68
prefer	41	0.15	38	0.00	+	186.87	5.24
ok	37	0.13	23	0.00	+	186.85	5.82
friends	54	0.19	111	0.01	+	185.00	4.10
earthquake	27	0.10	2	0.00	+	179.26	8.89
communication	31	0.11	10	0.00	+	177.41	6.77

answer	58	0.21	156	0.02	+	174.79	3.71
checklist	24	0.09	0	0.00	+	172.18	-9.19
tourism	25	0.09	1	0.00	+	170.93	9.78
student	35	0.13	28	0.00	+	166.11	5.46
information	67	0.24	262	0.03	+	162.81	3.17
city	27	0.10	6	0.00	+	162.75	7.30
teen	22	0.08	0	0.00	+	157.83	-9.31
p.	22	0.08	0	0.00	+	157.83	-9.31
activities	37	0.13	43	0.00	+	157.40	4.92
short	43	0.15	76	0.01	+	157.06	4.31
match	35	0.13	37	0.00	+	153.41	5.05
invitation	27	0.10	10	0.00	+	151.08	6.57
are	306	1.09	4812	0.49	+	148.14	1.16
Dan	24	0.09	5	0.00	+	145.80	7.40
follow	38	0.14	60	0.01	+	145.11	4.48
email	20	0.07	0	0.00	+	143.48	-9.45
adventures	20	0.07	0	0.00	+	143.48	-9.45
natural	32	0.11	32	0.00	+	142.65	5.13
study	32	0.11	34	0.00	+	140.05	5.05
can	254	0.91	3758	0.38	+	139.70	1.25
paragraph	34	0.12	47	0.00	+	136.37	4.67
modem	19	0.07	0	0.00	+	136.31	-9.52
about	230	0.82	3284	0.33	+	135.65	1.30
movie	23	0.08	7	0.00	+	132.80	6.85
texts	18	0.06	0	0.00	+	129.13	-9.60
rafting	18	0.06	0	0.00	+	129.13	-9.60
noun	18	0.06	0	0.00	+	129.13	-9.60
dishes	29	0.10	30	0.00	+	127.96	5.09
friendship	23	0.08	9	0.00	+	127.49	6.49
boxes	29	0.10	31	0.00	+	126.68	5.04
school	64	0.23	348	0.04	+	122.83	2.69
listen	44	0.16	143	0.01	+	119.64	3.43
daily	21	0.08	7	0.00	+	119.56	6.72
is	490	1.75	9985	1.02	+	118.07	0.79
teens	18	0.06	2	0.00	+	116.24	8.30
hi	28	0.10	36	0.00	+	115.18	4.77
make	83	0.30	645	0.07	+	115.05	2.18
messages	23	0.08	15	0.00	+	114.87	5.75
extreme_sports	16	0.06	0	0.00	+	114.79	-9.77
bike	31	0.11	53	0.01	+	114.75	4.36
chart	18	0.06	3	0.00	+	112.08	7.72
brochure	18	0.06	3	0.00	+	112.08	7.72

forces	19	0.07	5	0.00	+	112.03	7.06
cycling	17	0.06	2	0.00	+	109.29	8.22
mom	15	0.05	0	0.00	+	107.61	-9.86
tv	16	0.06	1	0.00	+	107.24	9.13
class	29	0.10	51	0.01	+	106.14	4.32
help	53	0.19	282	0.03	+	103.48	2.72
ingredients	17	0.06	4	0.00	+	101.73	7.22
tourist attractions	14	0.05	0	0.00	+	100.44	-9.96
parkour	14	0.05	0	0.00	+	100.44	-9.96
dossier	14	0.05	0	0.00	+	100.44	-9.96
understand	45	0.16	200	0.02	+	100.37	2.98
connection	20	0.07	13	0.00	+	99.96	5.76
pairs	18	0.06	7	0.00	+	99.88	6.50
with	239	0.85	4031	0.41	+	98.52	1.06
phrases	16	0.06	3	0.00	+	98.38	7.55
Sam	16	0.06	3	0.00	+	98.38	7.55
complete	29	0.10	61	0.01	+	98.34	4.06
in	489	1.75	10563	1.07	+	95.71	0.70
statements	20	0.07	16	0.00	+	94.92	5.46
city	27	0.10	54	0.01	+	93.62	4.13
recording	22	0.08	26	0.00	+	93.08	4.89
hello	41	0.15	180	0.02	+	92.23	3.00
world	42	0.15	196	0.02	+	90.50	2.91
box	33	0.12	109	0.01	+	88.89	3.41
message	23	0.08	37	0.00	+	87.20	4.45
responsibilities	17	0.06	10	0.00	+	86.93	5.90
false	15	0.05	5	0.00	+	85.40	6.72
spoken	21	0.08	28	0.00	+	85.30	4.72
expressions	14	0.05	3	0.00	+	84.76	7.36
fun	22	0.08	34	0.00	+	84.70	4.51
historic	16	0.06	8	0.00	+	84.68	6.13
correct	28	0.10	77	0.01	+	83.42	3.68
popular	19	0.07	20	0.00	+	83.39	5.06
water	49	0.18	311	0.03	+	82.55	2.47
net	18	0.06	17	0.00	+	81.60	5.22
Helen	18	0.06	18	0.00	+	80.24	5.13
vacation	12	0.04	1	0.00	+	79.09	8.72
website	11	0.04	0	0.00	+	78.92	-10.31
dna	11	0.04	0	0.00	+	78.92	-10.31
Onda	11	0.04	0	0.00	+	78.92	-10.31
ask	55	0.20	415	0.04	+	78.58	2.22
often	35	0.13	156	0.02	+	77.91	2.98

Bob	20	0.07	30	0.00	+	77.87	4.55
disasters	15	0.05	9	0.00	+	76.36	5.87
Kate	14	0.05	6	0.00	+	76.34	6.36
cable	17	0.06	17	0.00	+	75.78	5.13
words	37	0.13	185	0.02	+	75.78	2.81
brother	29	0.10	104	0.01	+	74.39	3.29
Frank	16	0.06	14	0.00	+	74.12	5.33
visit	22	0.08	48	0.00	+	73.38	4.01
Frank	16	0.06	15	0.00	+	72.69	5.23
habits	13	0.05	5	0.00	+	72.27	6.51
m.	11	0.04	1	0.00	+	72.09	8.59
exchange	17	0.06	20	0.00	+	72.03	4.90
mosque	10	0.04	0	0.00	+	71.74	-10.45
kayaking	10	0.04	0	0.00	+	71.74	-10.45
emails	10	0.04	0	0.00	+	71.74	-10.45
dislikes	10	0.04	0	0.00	+	71.74	-10.45
cooking	19	0.07	31	0.00	+	71.64	4.43
interaction	12	0.04	3	0.00	+	71.25	7.13
parents	24	0.09	70	0.01	+	69.31	3.59
computer	25	0.09	79	0.01	+	69.07	3.47
scan	12	0.04	4	0.00	+	68.32	6.72
cycle	11	0.04	2	0.00	+	67.87	7.59
enjoy	22	0.08	57	0.01	+	67.57	3.76
homework	18	0.06	31	0.00	+	66.44	4.35
Turkey	13	0.05	8	0.00	+	65.80	5.83
responsible	20	0.07	45	0.00	+	65.77	3.96
vegetable	10	0.04	1	0.00	+	65.10	8.46
weekdays	9	0.03	0	0.00	+	64.57	-10.60
web	9	0.03	0	0.00	+	64.57	-10.60
soccer	9	0.03	0	0.00	+	64.57	-10.60
sci-fi	9	0.03	0	0.00	+	64.57	-10.60
pc	9	0.03	0	0.00	+	64.57	-10.60
invitations	9	0.03	0	0.00	+	64.57	-10.60
preferences	11	0.04	3	0.00	+	64.54	7.01
excuses	11	0.04	3	0.00	+	64.54	7.01
refuse	13	0.05	9	0.00	+	64.00	5.67
express	15	0.05	19	0.00	+	62.02	4.79
music	25	0.09	95	0.01	+	61.87	3.21
salt	16	0.06	25	0.00	+	61.34	4.49
pan	16	0.06	26	0.00	+	60.43	4.43
friend	24	0.09	90	0.01	+	59.89	3.23
s	59	0.21	600	0.06	+	59.65	1.79

going to	80	0.29	1003	0.10	+	59.43	1.49
hang	13	0.05	12	0.00	+	59.32	5.25
dangerous	18	0.06	41	0.00	+	58.85	3.95
routines	9	0.03	1	0.00	+	58.12	8.30
invention	9	0.03	1	0.00	+	58.12	8.30
bake	9	0.03	1	0.00	+	58.12	8.30
adventure	9	0.03	1	0.00	+	58.12	8.30
poster	13	0.05	13	0.00	+	57.95	5.13
tourist	10	0.04	3	0.00	+	57.86	6.87
entertaining	10	0.04	3	0.00	+	57.86	6.87
reasons	22	0.08	76	0.01	+	57.72	3.35
calling	14	0.05	18	0.00	+	57.59	4.77
tsunami	8	0.03	0	0.00	+	57.39	-10.77
mardin	8	0.03	0	0.00	+	57.39	-10.77
landslide	8	0.03	0	0.00	+	57.39	-10.77
eye_catching	8	0.03	0	0.00	+	57.39	-10.77
flour	11	0.04	6	0.00	+	57.18	6.01
answers	16	0.06	30	0.00	+	57.03	4.23
conversation	20	0.07	60	0.01	+	56.88	3.55
dish	13	0.05	14	0.00	+	56.66	5.03
search	15	0.05	25	0.00	+	56.09	4.40
sometimes	31	0.11	183	0.02	+	55.61	2.57
family	28	0.10	149	0.02	+	54.66	2.72
turn_off	14	0.05	21	0.00	+	54.51	4.55
describe	14	0.05	21	0.00	+	54.51	4.55
get_up	20	0.07	65	0.01	+	54.38	3.43
vegetables	13	0.05	16	0.00	+	54.27	4.83
parachute	9	0.03	2	0.00	+	54.25	7.30
dictionary	13	0.05	17	0.00	+	53.16	4.75
bulb	10	0.04	5	0.00	+	52.93	6.13
Sam	16	0.06	37	0.00	+	51.94	3.93
such_as	18	0.06	53	0.01	+	51.71	3.58
points	22	0.08	91	0.01	+	51.53	3.09
teenagers	9	0.03	3	0.00	+	51.24	6.72
scientific	9	0.03	3	0.00	+	51.24	6.72
networking	8	0.03	1	0.00	+	51.17	8.13
gliding	8	0.03	1	0.00	+	51.17	8.13
gale	8	0.03	1	0.00	+	51.17	8.13
Jamaica	8	0.03	1	0.00	+	51.17	8.13
challenging	10	0.04	6	0.00	+	50.91	5.87
thrillers	7	0.03	0	0.00	+	50.22	-10.96
tablespoons	7	0.03	0	0.00	+	50.22	-10.96

Word	Freq	%	Freq2	%2	+/-	Keyness	Value
inventions7		0.03	0	0.00	+	50.22	-10.96
example_example70.03			0	0.00	+	50.22	-10.96
destination7		0.03	0	0.00	+	50.22	-10.96
comparisons7		0.03	0	0.00	+	50.22	-10.96
Tarsus	7	0.03	0	0.00	+	50.22	-10.96
Curtis	7	0.03	0	0.00	+	50.22	-10.96
's	3	0.01	20857	2.12	-	1133.33	99.49
that	80	0.29	21117	2.15	-	706.99	-86.69
yeah	5	0.02	9494	0.97	-	483.36	-98.15
it	347	1.24	26228	2.67	-	261.54	-53.52
oh	6	0.02	5533	0.56	-	259.73	-96.19
got	2	0.01	4965	0.51	-	257.81	-98.58
was	38	0.14	8052	0.82	-	241.41	-83.42
we	96	0.34	11364	1.16	-	217.29	-70.32
no	19	0.07	5925	0.60	-	212.66	-88.74
I	514	1.84	31907	3.25	-	198.55	-43.41
know	6	0.02	3985	0.41	-	176.77	-94.71
right	4	0.01	3298	0.34	-	152.12	-95.74
just	15	0.05	4076	0.41	-	138.24	-87.07
they	108	0.39	9778	1.00	-	133.30	-61.20
so	42	0.15	5570	0.57	-	119.14	-73.51
get	14	0.05	3458	0.35	-	112.25	-85.78
he	87	0.31	7890	0.80	-	107.84	-61.27
well	48	0.17	5626	0.57	-	106.45	-70.03
yes	28	0.10	4297	0.44	-	104.05	-77.11
said	8	0.03	2758	0.28	-	102.72	-89.81
did	22	0.08	3841	0.39	-	102.19	-79.88
had	6	0.02	2554	0.26	-	101.77	-91.75
but	68	0.24	6561	0.67	-	98.03	-63.59
were	11	0.04	2867	0.29	-	95.44	-86.52
one	43	0.15	5026	0.51	-	94.81	-69.95
there	71	0.25	6504	0.66	-	90.24	-61.65
na	1	0.00	1774	0.18	-	89.80	-98.02
be	58	0.21	5700	0.58	-	87.32	-64.26
this	59	0.21	5416	0.55	-	75.37	-61.73
gon	1	0.00	1447	0.15	-	71.85	-97.57
if	50	0.18	4612	0.47	-	64.64	-61.92
think	31	0.11	3491	0.36	-	63.22	-68.81
all	35	0.13	3689	0.38	-	61.82	-66.67
you_know9		0.03	1975	0.20	-	60.36	-83.99
actually	2	0.01	1228	0.12	-	53.60	-94.28
want	9	0.03	1788	0.18	-	51.65	-82.32

| now | 28 | 0.10 | 2997 | 0.30 | - | 51.17 | -67.18 |

THE COMPARISON OF THE PARTS OF SPEECH OF THE WORDS IN THE

COURSEBOOK TO THE BRITISH NATIONAL CORPUS

Item	O1	%1	O2	%2		LL	Log Ratio
NN1	5188	18.55	84446	8.59	+	2325.21	1.11
NN2	1991	7.12	24407	2.48	+	1533.92	1.52
VV0	1811	6.47	26118	2.66	+	1047.66	1.28
APPGE	737	2.63	10384	1.06	+	445.81	1.32
AT	1762	6.30	38652	3.93	+	324.96	0.68
JJ	1554	5.55	36947	3.76	+	202.16	0.56
II	1570	5.61	40428	4.11	+	132.67	0.45
NNU	43	0.15	159	0.02	+	108.25	3.25
NP1	666	2.38	15229	1.55	+	103.49	0.62
IW	241	0.86	4230	0.43	+	89.95	1.00
FO	15	0.05	12	0.00	+	71.19	5.46
RGT	24	0.09	136	0.01	+	44.55	2.63
RRT	14	0.05	56	0.01	+	33.53	3.13
NNO	43	0.15	571	0.06	+	28.97	1.40
RRQ	192	0.69	4542	0.46	+	25.58	0.57
AT1	703	2.51	20274	2.06	+	24.89	0.28
NNT2	78	0.28	1463	0.15	+	24.31	0.91
DA2	59	0.21	1017	0.10	+	23.05	1.03
RGR	28	0.10	356	0.04	+	20.32	1.47
CC	949	3.39	28666	2.92	+	19.98	0.22
CSN	41	0.15	653	0.07	+	19.30	1.14
NPD1	50	0.18	889	0.09	+	18.05	0.98
JJT	30	0.11	476	0.05	+	14.24	1.15
NNT1	153	0.55	3920	0.40	+	13.33	0.46
ND1	12	0.04	129	0.01	+	11.25	1.71
VVG	445	1.59	13349	1.36	+	10.22	0.23
RGQ	39	0.14	802	0.08	+	9.11	0.77
VDG	48	0.17	1063	0.11	+	8.52	0.67
RR22	11	0.04	136	0.01	+	8.36	1.51
VD0	264	0.94	7738	0.79	+	7.91	0.26
VDI	95	0.34	2464	0.25	+	7.68	0.44
NPD2	5	0.02	36	0.00	+	7.49	2.29
NN	108	0.39	2889	0.29	+	7.11	0.39
IF	209	0.75	6141	0.62	+	6.10	0.26
NN121	3	0.01	19	0.00	+	5.06	2.47

VBR	306	1.09	9496	0.97	+	4.38	0.18
NNU2	11	0.04	227	0.02	+	2.54	0.77
NNL2	1	0.00	4	0.00	+	2.39	3.13
RA	6	0.02	104	0.01	+	2.31	1.02
RG22	1	0.00	5	0.00	+	2.05	2.81
VDZ	53	0.19	1694	0.17	+	0.45	0.14
NNL1	4	0.01	104	0.01	+	0.32	0.43
JJ22	1	0.00	23	0.00	+	0.15	0.61
VBM	94	0.34	3173	0.32	+	0.14	0.06
PN	8	0.03	264	0.03	+	0.03	0.09
RRQV	2	0.01	63	0.01	+	0.02	0.16
PPX221	2	0.01	66	0.01	+	0.01	0.09
VVGK	66	0.24	2302	0.23	+	0.00	0.01
UH	140	0.50	33336	3.39	-	1062.87	85.25
DD1	109	0.39	21326	2.17	-	610.40	-82.05
XX	135	0.48	18780	1.91	-	419.31	-74.75
RR	598	2.14	39925	4.06	-	301.98	-47.38
PPH1	342	1.22	26247	2.67	-	269.70	-54.23
VVN	169	0.60	16468	1.68	-	249.40	-63.95
VBDZ	38	0.14	8118	0.83	-	244.49	-83.56
PPIS2	96	0.34	11292	1.15	-	214.46	-70.14
MC	108	0.39	11888	1.21	-	209.76	-68.09
PPIS1	529	1.89	32828	3.34	-	204.10	-43.39
VH0	90	0.32	9541	0.97	-	161.01	-66.86
PPHS1	155	0.55	12947	1.32	-	155.17	-57.94
RP	26	0.09	5154	0.52	-	148.69	-82.28
CST	51	0.18	6644	0.68	-	139.77	-73.04
PPHS2	109	0.39	9853	1.00	-	134.01	-61.14
VBZ	490	1.75	27564	2.80	-	124.83	-37.55
VHD	6	0.02	2921	0.30	-	120.76	-92.78
VVD	211	0.75	14200	1.44	-	109.54	-47.80
VDD	22	0.08	3853	0.39	-	102.72	-79.94
CCB	68	0.24	6516	0.66	-	96.43	-63.34
VBDR	11	0.04	2883	0.29	-	96.21	-86.60
PN1	74	0.26	6693	0.68	-	91.11	-61.16
VBI	60	0.21	5691	0.58	-	83.03	-62.96
GE	3	0.01	1883	0.19	-	82.58	-94.40
CS	196	0.70	12317	1.25	-	79.37	-44.10
EX	32	0.11	3812	0.39	-	73.38	-70.51
VHZ	31	0.11	3607	0.37	-	67.71	-69.81
DB	43	0.15	4162	0.42	-	62.46	-63.71
RG	85	0.30	6250	0.64	-	58.91	-52.23
VHI	40	0.14	3773	0.38	-	54.62	-62.76

VM	433	1.55	21085	2.15	-	50.35	-27.86
VDN	2	0.01	996	0.10	-	41.42	-92.95
MC1	35	0.13	2957	0.30	-	36.13	-58.42
PPY	590	2.11	26362	2.68	-	36.10	-21.38
PPHO2	46	0.16	3514	0.36	-	35.78	-54.02
VHN	1	0.00	729	0.07	-	32.91	-95.18
RR21	39	0.14	3052	0.31	-	32.55	-55.11
PPHO1	31	0.11	2481	0.25	-	27.58	-56.11
RT	182	0.65	9174	0.93	-	26.20	-30.31
CSW	3	0.01	726	0.07	-	23.33	-85.48
PPGE	1	0.00	528	0.05	-	22.28	-93.35
DD2	32	0.11	2311	0.24	-	20.97	-51.36
PPIO1	39	0.14	2640	0.27	-	20.66	-48.11
DA	8	0.03	863	0.09	-	14.87	-67.44
VBG	4	0.01	574	0.06	-	13.16	-75.52
DD	75	0.27	3857	0.39	-	12.12	-31.69
JK	1	0.00	301	0.03	-	10.65	-88.33
ZZ1	113	0.40	5328	0.54	-	10.55	-25.50
TO	309	1.10	12929	1.32	-	9.76	-16.04
DDQV	1	0.00	266	0.03	-	8.94	-86.79
NNB	15	0.05	954	0.10	-	6.36	-44.77
II31	8	0.03	570	0.06	-	5.02	-50.70
DA1	7	0.03	508	0.05	-	4.66	-51.59
CSA	26	0.09	1335	0.14	-	4.16	-31.59
PNQS	38	0.14	1827	0.19	-	4.06	-26.94
JJR	27	0.10	1280	0.13	-	2.62	-25.90
DAR	17	0.06	866	0.09	-	2.60	-31.04
RL	152	0.54	6020	0.61	-	2.22	-11.30
DB2	3	0.01	214	0.02	-	1.89	-50.75
DDQ	256	0.92	9728	0.99	-	1.58	-7.56
JJ21	1	0.00	99	0.01	-	1.53	-64.52
NNU1	4	0.01	245	0.02	-	1.47	-42.65
MD	69	0.25	2794	0.28	-	1.42	-13.25
PPX2	3	0.01	195	0.02	-	1.38	-45.96
CS31	4	0.01	241	0.02	-	1.37	-41.70
VVZ	173	0.62	6603	0.67	-	1.19	-7.96
IO	344	1.23	12770	1.30	-	1.04	-5.37
FW	1	0.00	83	0.01	-	0.98	-57.68
PPIO2	24	0.09	1025	0.10	-	0.95	-17.75
DAT	5	0.02	264	0.03	-	0.93	-33.47
NPM1	8	0.03	388	0.04	-	0.91	-27.57
VVI	946	3.38	34201	3.48	-	0.77	-2.84
VHG	12	0.04	502	0.05	-	0.38	-16.03

REX21	3	0.01	146	0.01	-	0.35	-27.82
RG21	1	0.00	59	0.01	-	0.31	-40.46
II21	40	0.14	1486	0.15	-	0.12	-5.44
NNO2	2	0.01	84	0.01	-	0.07	-16.36
RRR	23	0.08	844	0.09	-	0.04	-4.27
NP2	2	0.01	79	0.01	-	0.03	-11.07
PPX1	13	0.05	467	0.05	-	0.01	-2.21

APPENDIX G

THE COMPARISON OF THE SEMANTIC FIELDS OF THE WORDS IN THE COURSEBOOK

TO THE BRITISH NATIONAL CORPUS

Item	O1	%1	O2	%2		LL	Log Ratio	
Q3	406	1.45	1105	0.11	+	1216.06	3.69	Language, speech and grammar
Y2	284	1.02	480	0.05	+	1056.13	4.38	Information technology and computing
Q1.2	465	1.66	3398	0.35	+	686.14	2.27	Paper documents and writing
F1	474	1.69	3914	0.40	+	615.73	2.09	Food
P1	323	1.15	1928	0.20	+	574.03	2.56	Education in general
S3.1	174	0.62	752	0.08	+	395.69	3.02	Personal relationship: General
Q2.2	478	1.71	5794	0.59	+	374.93	1.54	Speech acts
O2	392	1.40	4156	0.42	+	374.66	1.73	Objects generally
K5.1	169	0.60	914	0.09	+	325.74	2.70	Sports
Y1	116	0.41	415	0.04	+	298.01	3.30	Science and technology in general
W3	142	0.51	699	0.07	+	294.25	2.84	Geographical terms
S5+	250	0.89	2602	0.26	+	245.04	1.75	Belonging to a group
E2++	57	0.20	76	0.01	+	231.54	4.72	Like
A2.2	200	0.71	1891	0.19	+	221.93	1.89	Cause Effect/Connection
X2.4	135	0.48	889	0.09	+	219.98	2.42	Investigate, examine, test, search
A1.5.1	153	0.55	1213	0.12	+	207.66	2.15	Using
A6.2+	136	0.49	1087	0.11	+	183.00	2.14	Comparing: Usual
Q4.3	98	0.35	562	0.06	+	180.12	2.61	The Media: TV, Radio and Cinema
E2+	200	0.71	2255	0.23	+	175.15	1.64	Like

W4	77	0.28	379	0.04	+	159.57	2.84	Weather
Z1	499	1.78	9430	0.96	+	152.17	0.89	Personal names
C1	110	0.39	896	0.09	+	145.03	2.11	Arts and crafts
S1.1.2+	80	0.29	471	0.05	+	143.84	2.58	Reciprocal
Z2	240	0.86	3541	0.36	+	132.74	1.25	Geographical names
A12+	65	0.23	347	0.04	+	126.58	2.72	Easy
E2+++	32	0.11	47	0.00	+	125.56	4.58	Like
B4	117	0.42	1152	0.12	+	123.37	1.83	Cleaning and personal care
W1	60	0.21	300	0.03	+	122.89	2.81	The universe
A6.1	21	0.08	9	0.00	+	114.51	6.36	Comparing: Similar/different
T1.3-	44	0.16	158	0.02	+	112.78	3.29	Time period: short
A15-	41	0.15	140	0.01	+	108.32	3.36	Danger
A1.1.1	645	2.31	14522	1.48	+	107.06	0.64	General actions / making
A9	18	0.06	5	0.00	+	105.33	6.98	Getting and giving; possession
S1.1.1	85	0.30	746	0.08	+	103.09	2.00	Social Actions, States and Processes
A5.1---	21	0.08	29	0.00	+	84.26	4.67	Evaluation: Bad
Q1.3	79	0.28	783	0.08	+	82.60	1.83	Telecommunications
O1.2	75	0.27	732	0.07	+	79.98	1.85	Substances and materials: Liquid
K2	66	0.24	586	0.06	+	78.98	1.98	Music and related activities
K1	97	0.35	1201	0.12	+	73.54	1.50	Entertainment generally
X5.2+	62	0.22	612	0.06	+	65.17	1.83	Interested/excited/energetic
X9.2++	9	0.03	0	0.00	+	64.57	-10.60	Success
X3.2	88	0.31	1139	0.12	+	61.98	1.44	Sensory: Sound
O3	62	0.22	646	0.07	+	60.68	1.75	Electricity and electrical equipment
S4	191	0.68	3699	0.38	+	54.15	0.86	Kin
Q4.1	58	0.21	635	0.06	+	53.00	1.68	The Media: Books
W2	7	0.03	0	0.00	+	50.22	-10.96	Light

N6+	107	0.38	1719	0.17	+	49.40	1.13	Frequent
M3	122	0.44	2115	0.22	+	46.99	1.02	Vehicles and transport on land
S8+	118	0.42	2020	0.21	+	46.91	1.04	Helping
N6	42	0.15	399	0.04	+	46.33	1.89	Frequency
W5	20	0.07	88	0.01	+	44.92	3.00	Green issues
S2	144	0.51	2728	0.28	+	43.60	0.89	People
X3.1+	6	0.02	0	0.00	+	43.04	-11.19	Tasty
M5	32	0.11	269	0.03	+	40.76	2.06	Flying and aircraft
M7	161	0.58	3261	0.33	+	39.59	0.79	Places
X2.5+	53	0.19	684	0.07	+	37.51	1.44	Understanding
K5.2	20	0.07	115	0.01	+	36.68	2.61	Games
H4	72	0.26	1130	0.11	+	35.01	1.16	Residence
S1.2.4+	18	0.06	103	0.01	+	33.14	2.62	Polite
L1+	13	0.05	51	0.01	+	31.52	3.16	Alive
A4.1	119	0.43	2356	0.24	+	31.50	0.83	Generally kinds, groups, examples
A6.1+	60	0.21	925	0.09	+	30.31	1.19	Comparing: Similar
N5---	9	0.03	20	0.00	+	29.77	3.98	Quantities: little
T1.3	288	1.03	7171	0.73	+	29.56	0.50	Time: Period
N3.7+++	6	0.02	5	0.00	+	28.17	5.40	Long, tall and wide
M4	36	0.13	458	0.05	+	26.10	1.47	Sailing, swimming, etc.
A1.2+	21	0.08	195	0.02	+	23.82	1.92	Suitable
K4	25	0.09	269	0.03	+	23.41	1.71	Drama, the theatre and show business
N3	4	0.01	2	0.00	+	21.17	6.13	Measurement
X7-	13	0.05	88	0.01	+	20.65	2.38	Unwanted
O1	28	0.10	362	0.04	+	19.76	1.44	Substances and materials generally
O4.6+	50	0.18	863	0.09	+	19.47	1.03	Temperature: Hot / on fire
A5.2-	18	0.06	180	0.02	+	18.60	1.81	Evaluation: False

A5	6	0.02	16	0.00	+	18.16	3.72	Evaluation
L3	36	0.13	564	0.06	+	17.57	1.16	Plants
T3	47	0.17	868	0.09	+	15.31	0.93	Time: Old, new and young; age
A10+	79	0.28	1730	0.18	+	14.67	0.68	Open; Finding; Showing
A6.2--	2	0.01	0	0.00	+	14.35	-12.77	Comparing: Unusual
N6-	17	0.06	205	0.02	+	13.44	1.54	Infrequent
T4+	12	0.04	113	0.01	+	13.38	1.90	Time: Early
S1.1.4+	5	0.02	17	0.00	+	13.24	3.37	Deserving
S7.3	4	0.01	9	0.00	+	13.15	3.96	Competition
K5	8	0.03	60	0.01	+	11.50	2.23	Sports and games generally
N3.2+++	11	0.04	112	0.01	+	11.10	1.79	Size: Big
S1.2.1+	9	0.03	78	0.01	+	11.07	2.02	Informal/Friendly
X9.2+	33	0.12	603	0.06	+	11.07	0.94	Success
E4.1+	40	0.14	782	0.08	+	11.02	0.85	Happy
N3.3	18	0.06	260	0.03	+	10.38	1.28	Measurement: Distance
N3.6	8	0.03	67	0.01	+	10.23	2.07	Measurement: Area
Z99	204	0.73	5684	0.58	+	9.85	0.33	Unmatched
N3.7+	20	0.07	320	0.03	+	9.32	1.13	Long, tall and wide
M1	420	1.50	12692	1.29	+	8.79	0.22	Moving, coming and going
O4.6	6	0.02	45	0.00	+	8.63	2.23	Temperature
T2+++	5	0.02	31	0.00	+	8.60	2.50	Time: Beginning
X3.2+	4	0.01	20	0.00	+	8.19	2.81	Sound: Loud
S1.2.5+	10	0.04	119	0.01	+	8.07	1.56	Tough/strong
S7.3+	8	0.03	82	0.01	+	8.00	1.78	Competitive
A1.5	2	0.01	3	0.00	+	7.79	4.55	Use
N3.2	17	0.06	276	0.03	+	7.66	1.11	Measurement: Size
A1.1.2	27	0.10	524	0.05	+	7.61	0.86	Damaging and destroying

X2	3	0.01	12	0.00	+	7.18	3.13	Mental actions and processes
A1.9-	1	0.00	0	0.00	+	7.17	-13.77	Unavoidable
A13	1	0.00	0	0.00	+	7.17	-13.77	Degree
A5.2+	26	0.09	513	0.05	+	6.96	0.83	Evaluation: True
H4-	2	0.01	4	0.00	+	6.93	4.13	Non-resident
X2.6-	5	0.02	39	0.00	+	6.90	2.17	Unexpected
A11.1+	42	0.15	955	0.10	+	6.69	0.63	Important
H5	53	0.19	1276	0.13	+	6.49	0.55	Furniture and household fittings
H1	68	0.24	1725	0.18	+	6.28	0.47	Architecture, houses and buildings
X5.2++	2	0.01	5	0.00	+	6.25	3.81	Interested/excited/energetic
A1.8+	23	0.08	457	0.05	+	6.02	0.82	Inclusion
A12++	7	0.03	81	0.01	+	5.90	1.60	Easy
M2	188	0.67	5496	0.56	+	5.80	0.26	Putting, pulling, pushing, transporting
E2-	16	0.06	286	0.03	+	5.69	0.97	Dislike
T3-	62	0.22	1580	0.16	+	5.58	0.46	Time: New and young
O1.1	46	0.16	1117	0.11	+	5.39	0.53	Substances and materials: Solid
O4.6-	14	0.05	247	0.03	+	5.16	0.99	Temperature: Cold
A15++	2	0.01	8	0.00	+	4.79	3.13	Safe
Q1.1	95	0.34	2643	0.27	+	4.64	0.34	Linguistic Actions, States and Processes; Communication
X8+	48	0.17	1213	0.12	+	4.53	0.48	Trying hard
T3+	25	0.09	561	0.06	+	4.21	0.65	Time: Old; grown-up
N3.2-	44	0.16	1114	0.11	+	4.11	0.47	Size: Small
X4.1	42	0.15	1066	0.11	+	3.87	0.47	Mental object: Conceptual object
A5.4-	13	0.05	254	0.03	+	3.59	0.85	Evaluation: Unauthentic
O4.1	35	0.13	890	0.09	+	3.19	0.47	General appearance and physical properties
A5.1+++	19	0.07	428	0.04	+	3.15	0.64	Evaluation: Good
T4-	10	0.04	188	0.02	+	3.10	0.90	Time: Late

N5+++	14	0.05	294	0.03	+	3.04	0.74	Quantities: many/much
A1.6	5	0.02	71	0.01	+	2.98	1.31	Concrete/Abstract
F4	13	0.05	270	0.03	+	2.94	0.76	Farming & Horticulture
A6.3+	12	0.04	244	0.02	+	2.91	0.79	Comparing: Varied
G3-	1	0.00	3	0.00	+	2.84	3.55	Anti-war
O4.3	61	0.22	1718	0.17	+	2.69	0.32	Colour and colour patterns
Z5	7321	26.17	252254	25.67	+	2.64	0.03	Grammatical bin
N6+++	26	0.09	650	0.07	+	2.61	0.49	Frequent
T2-	52	0.19	1442	0.15	+	2.61	0.34	Time: Ending
X8+++	1	0.00	4	0.00	+	2.39	3.13	Trying hard
B2++	1	0.00	4	0.00	+	2.39	3.13	Healthy
B3	38	0.14	1024	0.10	+	2.38	0.38	Medicines and medical treatment
X2.6+	12	0.04	261	0.03	+	2.29	0.69	Expected
A11.1++	1	0.00	5	0.00	+	2.05	2.81	Important
O1.2-	6	0.02	109	0.01	+	2.04	0.95	Dry
A13.2	37	0.13	1015	0.10	+	2.02	0.36	Degree: Maximizers
H2	61	0.22	1774	0.18	+	1.99	0.27	Parts of buildings
Q4.2	18	0.06	447	0.05	+	1.87	0.50	The Media: Newspapers etc.
N3.8+	30	0.11	819	0.08	+	1.70	0.36	Speed: Fast
A2.1+	68	0.24	2031	0.21	+	1.65	0.23	Change
A15+	6	0.02	120	0.01	+	1.54	0.81	Safe
Q4	7	0.03	147	0.01	+	1.52	0.74	The Media
T4--	8	0.03	175	0.02	+	1.49	0.68	Time: Late
N3.3--	2	0.01	26	0.00	+	1.40	1.43	Distance: Near
K3	11	0.04	264	0.03	+	1.37	0.55	Recorded sound
A11.2+	8	0.03	184	0.02	+	1.21	0.61	Noticeable
A6.2-	10	0.04	242	0.02	+	1.19	0.54	Comparing: Unusual

N5.1-	31	0.11	908	0.09	+	0.94	0.26	Part
N3.8++++	1	0.00	12	0.00	+	0.80	1.55	Speed: Fast
N3.8-	4	0.01	90	0.01	+	0.67	0.64	Speed: Slow
Z3	82	0.29	2633	0.27	+	0.62	0.13	Other proper names
S1.2.5++	1	0.00	15	0.00	+	0.53	1.23	Tough/strong
F2	33	0.12	1022	0.10	+	0.49	0.18	Drinks and alcohol
X2.2-	10	0.04	281	0.03	+	0.45	0.32	No knowledge
A3-	1	0.00	17	0.00	+	0.40	1.05	Non-existing
A2.1-	3	0.01	72	0.01	+	0.37	0.55	No change
N3.7-	4	0.01	104	0.01	+	0.32	0.43	Short and narrow
X5.2-	6	0.02	170	0.02	+	0.25	0.31	Uninterested/bored/unenergetic
X6+	16	0.06	511	0.05	+	0.14	0.14	Decided
B2	5	0.02	148	0.02	+	0.13	0.25	Health and disease
S1.2.1-	2	0.01	55	0.01	+	0.11	0.35	Formal/Unfriendly
N3.3---	1	0.00	25	0.00	+	0.10	0.49	Distance: Near
N3.5+	3	0.01	88	0.01	+	0.09	0.26	Weight: Heavy
A1.5.1+	1	0.00	28	0.00	+	0.05	0.33	Used
B5	61	0.22	2080	0.21	+	0.05	0.04	Clothes and personal belongings
X3.3	2	0.01	69	0.01	+	0.00	0.03	Sensory: Touch
E3+	6	0.02	207	0.02	+	0.00	0.03	Calm
T1.1	1	0.00	35	0.00	+	0.00	0.01	Time: General
T1.3++	2	0.01	70	0.01	+	0.00	0.01	Time period: long
X5.1+	4	0.01	140	0.01	+	0.00	0.01	Attentive
Z4	294	1.05	65278	6.64	-	2007.66	-84.18	Discourse Bin
Z8	3109	11.11	172345	17.54	-	739.48	-36.63	Pronouns
Z6	148	0.53	19932	2.03	-	432.39	-73.92	Negative
N1	129	0.46	13063	1.33	-	208.17	-65.31	Numbers

A3+	788	2.82	43253	4.40	-	178.75	-36.00	Existing
A14	22	0.08	5480	0.56	-	178.60	-85.90	Exclusivizers/particularizers
A9+	275	0.98	18577	1.89	-	144.63	-48.00	Getting and possession
A5.4+	6	0.02	3324	0.34	-	141.83	-93.66	Evaluation: Authentic
Z7	50	0.18	4767	0.49	-	70.05	-63.15	If
S2.2	5	0.02	1829	0.19	-	69.51	-90.40	People: Male
N5.1+	81	0.29	6294	0.64	-	66.24	-54.79	Entire; maximum
I1	11	0.04	2197	0.22	-	63.67	-82.41	Money generally
T1.1.3	163	0.58	9716	0.99	-	53.50	-41.07	Time: Future
X7+	71	0.25	5302	0.54	-	51.60	-52.96	Wanted
A13.5	2	0.01	1189	0.12	-	51.54	-94.09	Degree: Compromisers
X2.1	113	0.40	7031	0.72	-	44.05	-43.54	Thought, belief
S7.1+	33	0.12	3064	0.31	-	43.36	-62.17	In power
N5-	12	0.04	1766	0.18	-	41.35	-76.13	Quantities: little
T1.1.1	25	0.09	2465	0.25	-	37.93	-64.37	Time: Past
T2++	12	0.04	1678	0.17	-	37.63	-74.88	Time: Beginning
A5.1+	191	0.68	10112	1.03	-	36.12	-33.65	Evaluation: Good
X4.2	10	0.04	1464	0.15	-	34.13	-76.01	Mental object: Means, method
S6+	93	0.33	5698	0.58	-	34.12	-42.67	Strong obligation or necessity
I2.1	6	0.02	1086	0.11	-	29.60	-80.59	Business: Generally
G2.1	15	0.05	1639	0.17	-	28.67	-67.85	Law and order
S2.1	4	0.01	875	0.09	-	26.70	-83.94	People: Female
A7	3	0.01	784	0.08	-	26.13	-86.56	Probability
A6.1-	61	0.22	3792	0.39	-	23.69	-43.49	Comparing: Different
G1.2	1	0.00	554	0.06	-	23.64	-93.66	Politics
A8	15	0.05	1482	0.15	-	22.87	-64.45	Seem
T1.3+	4	0.01	709	0.07	-	19.06	-80.18	Time period: long

T1.1.2	90	0.32	4869	0.50	-	19.01	-35.07	Time: Present; simultaneous
A5.1	1	0.00	418	0.04	-	16.57	-91.60	Evaluation: Good/bad
H3	2	0.01	470	0.05	-	14.89	-85.05	Areas around or near houses
E3-	10	0.04	959	0.10	-	14.21	-63.37	Violent/Angry
X9.1+	8	0.03	842	0.09	-	14.08	-66.62	Able/intelligent
I2.2	50	0.18	2864	0.29	-	13.83	-38.67	Business: Selling
A6.1+++	6	0.02	648	0.07	-	11.18	-67.47	Comparing: Similar
A1.4	2	0.01	379	0.04	-	10.64	-81.46	Chance, luck
N4	137	0.49	6284	0.64	-	10.41	-23.42	Linear order
A9-	53	0.19	2788	0.28	-	9.69	-33.22	Giving
I1.2	16	0.06	1117	0.11	-	9.40	-49.68	Money: Debts
M6	278	0.99	11645	1.18	-	8.90	-16.14	Location and direction
S1.2.6-	1	0.00	264	0.03	-	8.84	-86.69	Foolish
G2.2-	2	0.01	338	0.03	-	8.79	-79.21	Unethical
G1.1	16	0.06	1084	0.11	-	8.50	-48.15	Government
A13.3	121	0.43	5457	0.56	-	8.01	-22.11	Degree: Boosters
I4	1	0.00	245	0.02	-	7.92	-85.66	Industry
I1.1	20	0.07	1235	0.13	-	7.57	-43.11	Money and pay
A1.9	1	0.00	229	0.02	-	7.16	-84.66	Avoiding
E6-	10	0.04	733	0.07	-	6.86	-52.08	Worry
A5.1++	6	0.02	507	0.05	-	6.20	-58.43	Evaluation: Good
I3.1-	1	0.00	200	0.02	-	5.80	-82.44	Unemployed
S7.4+	15	0.05	930	0.09	-	5.77	-43.34	Allowed
I3.1	58	0.21	2747	0.28	-	5.59	-25.83	Work and employment: Generally
A13.1	6	0.02	487	0.05	-	5.55	-56.72	Degree: Non-specific
X2.2+	128	0.46	5481	0.56	-	5.23	-17.96	Knowledgeable
A1.3+	1	0.00	181	0.02	-	4.93	-80.59	Cautious

A13.6	18	0.06	1022	0.10	-	4.78	-38.13	Degree: Diminishers
I1.3	17	0.06	976	0.10	-	4.75	-38.81	Money: Cost and price
N3.3-	4	0.01	355	0.04	-	4.69	-60.42	Distance:Near
N5++	54	0.19	2514	0.26	-	4.60	-24.55	Quantities: many/much
A5.3-	5	0.02	401	0.04	-	4.48	-56.20	Evaluation: Inaccurate
E4.1-	5	0.02	400	0.04	-	4.44	-56.09	Sad
G3	21	0.08	1128	0.11	-	4.28	-34.60	Warfare, defence and the army; weapons
G2.2+	2	0.01	225	0.02	-	4.07	-68.77	Ethical
A5.1-	11	0.04	656	0.07	-	3.62	-41.10	Evaluation: Bad
N3.3+	1	0.00	148	0.02	-	3.48	-76.26	Distance:Far
N3.4	2	0.01	205	0.02	-	3.32	-65.73	Measurement: Volume
T3---	1	0.00	143	0.01	-	3.27	-75.43	Time: New and young
S9	22	0.08	1106	0.11	-	3.12	-30.12	Religion and the supernatural
O4.2+	37	0.13	1717	0.17	-	3.07	-24.30	Judgement of appearance: Beautiful
A4.2+	22	0.08	1082	0.11	-	2.73	-28.57	Detailed
G2.1-	4	0.01	290	0.03	-	2.65	-51.55	Crime
X3.4	96	0.34	3968	0.40	-	2.61	-15.01	Sensory: Sight
I1.3-	3	0.01	237	0.02	-	2.57	-55.53	Cheap
T1	53	0.19	2290	0.23	-	2.37	-18.70	Time
A13.4	30	0.11	1380	0.14	-	2.33	-23.63	Degree: Approximators
X3.2-	1	0.00	115	0.01	-	2.13	-69.45	Sound: Quiet
O4.2-	12	0.04	628	0.06	-	2.13	-32.88	Judgement of appearance: Ugly
A7+	398	1.42	15034	1.53	-	2.09	-7.00	Likely
O4.4	25	0.09	1156	0.12	-	2.02	-24.03	Shape
M8	17	0.06	822	0.08	-	1.89	-27.35	Stationary
N3.2++	1	0.00	108	0.01	-	1.86	-67.47	Size: Big
I1.1+	1	0.00	99	0.01	-	1.53	-64.52	Money: Affluence

S3.2	11	0.04	549	0.06	-	1.49	-29.62	Relationship: Intimacy and sex
N3.5	3	0.01	195	0.02	-	1.38	-45.96	Measurement: Weight
S1.1.3+	14	0.05	660	0.07	-	1.31	-25.49	Participating
N3.2+	24	0.09	1046	0.11	-	1.17	-19.40	Size: Big
A10-	15	0.05	683	0.07	-	1.07	-22.85	Closed; Hiding/Hidden
A7+++	2	0.01	135	0.01	-	1.05	-47.96	Likely
E4.2+	3	0.01	178	0.02	-	0.97	-40.80	Content
A5.3+	29	0.10	1206	0.12	-	0.85	-15.53	Evaluation: Accurate
S7.1-	4	0.01	214	0.02	-	0.80	-34.34	No power
N5--	1	0.00	75	0.01	-	0.74	-53.16	Quantities: little
S1.2.4-	1	0.00	74	0.01	-	0.71	-52.53	Impolite
N3.7	6	0.02	292	0.03	-	0.70	-27.82	Measurement: Length & height
X9.2-	4	0.01	207	0.02	-	0.67	-32.12	Failure
O4.5	10	0.04	452	0.05	-	0.67	-22.28	Texture
N5	141	0.50	5292	0.54	-	0.61	-6.40	Quantities
N3.8++	1	0.00	70	0.01	-	0.59	-49.82	Speed: Fast
L2	44	0.16	1727	0.18	-	0.55	-10.50	Living creatures: animals, birds, etc.
O1.3	7	0.03	320	0.03	-	0.52	-23.16	Substances and materials: Gas
T2+	35	0.13	1386	0.14	-	0.51	-11.29	Time: Beginning
E5-	9	0.03	398	0.04	-	0.50	-20.56	Fear/shock
K6	2	0.01	110	0.01	-	0.46	-36.13	Children's games and toys
X1	1	0.00	63	0.01	-	0.41	-44.24	Psychological Actions, States and Processes
B2+	2	0.01	106	0.01	-	0.38	-33.72	Healthy
L1-	12	0.04	501	0.05	-	0.37	-15.86	Dead
A13.7	11	0.04	462	0.05	-	0.36	-16.36	Degree: Minimizers
X2.3+	8	0.03	341	0.03	-	0.31	-17.59	Learning
T1.2	26	0.09	1015	0.10	-	0.29	-10.02	Time: Momentary

T3+++	1	0.00	57	0.01	-	0.27	-38.37	Time: Old; grown-up
A1.7+	10	0.04	413	0.04	-	0.27	-14.94	Constraint
A11.1+++	1	0.00	56	0.01	-	0.25	-37.27	Important
X2.5-	2	0.01	95	0.01	-	0.20	-26.05	Not understanding
S7.2+	2	0.01	94	0.01	-	0.18	-25.26	Respected
A1.7-	6	0.02	250	0.03	-	0.18	-15.69	No constraint
S1.1.3-	1	0.00	51	0.01	-	0.15	-31.12	Non-participating
G2.2	1	0.00	51	0.01	-	0.15	-31.12	General ethics
N3.2--	1	0.00	50	0.01	-	0.14	-29.74	Size: Small
N5+	100	0.36	3627	0.37	-	0.10	-3.15	Quantities: many/much
F1-	1	0.00	47	0.00	-	0.09	-25.26	Lack of food
N3.1	8	0.03	308	0.03	-	0.07	-8.76	Measurement: General
S5-	9	0.03	340	0.03	-	0.05	-7.01	Not part of a group
N2	11	0.04	413	0.04	-	0.05	-6.44	Mathematics
N5.2+	21	0.08	771	0.08	-	0.04	-4.32	Exceed; waste
A12-	29	0.10	1056	0.11	-	0.04	-3.53	Difficult
Q2.1	285	1.02	10107	1.03	-	0.03	-0.94	Speech: Communicative
B1	104	0.37	3703	0.38	-	0.02	-1.34	Anatomy and physiology
T3++	3	0.01	108	0.01	-	0.00	-2.42	Time: Old; grown-up
B2-	27	0.10	959	0.10	-	0.00	-1.10	Disease
S1.2.5-	2	0.01	71	0.01	-	0.00	-1.05	Weak
S6-	3	0.01	106	0.01	-	0.00	-0.58	No obligation or necessity

APPENDIX H

QUESTIONNAIRE

Değerli Öğrenciler,

Bu çalışma, Ondokuz Mayıs Üniversitesi Eğitim Bilimleri Enstitüsü Yabancı Diller Eğitimi Anabilim Dalı Yüksek Lisans programı için hazırlanan *"İlköğretim A2 Seviyesindeki İngilizce Ders Kitaplarındaki Kelimeler Üzerinde Derlem Tabanlı Çalışma"* başlıklı tez konusu için hazırlanmıştır. Vereceğiniz bilgiler yalnızca araştırma amaçlı kullanılacak olup vereceğiniz cevaplar saklı kalacaktır. İçten yanıtlarınızla katkıda bulunduğunuz için teşekkür ederiz.

İngilizce Öğretmeni
Doğan BAYDAL

8. sınıf İngilizce ders kitabında yer almasını istediğiniz, ilgi duyduğunuz alanları aşağıda belirtiniz.

CURRICULUM VITAE

Education:
High School: Çarşamba High School (Foreign Language Intensive)/2006
Bachelor of Arts: University College of Lillebaelt, Odense Seminarium, English Language Teaching (Erasmus Student)/2008
Bachelor of Arts: Erciyes University, Faculty of Education, English Language Teaching/2010

Work Experience:
2010-… English Language Teacher / Ministry of National Education

Awards:
Certificate of Honour: (Erciyes University, 2006-2007, 2007-2008, 2009-2010 Academic Years)

Certificates:
Certificate of "The Regional English Language Office of the Embassy of the United States of America, 2014 Summer English Language Teaching Institute, 2014, Trabzon/TURKEY

Amasya Üniversitesi, Türkiye'de İngilizce Öğretmeni Yetiştirme Programlarının İşleyişi, Sorunları ve Çözüm Önerileri Çalıştayı, Katılım Sertifikası, 2014, Amasya/TURKEY

MEB, Aktif Öğretim Yöntem ve Teknikleri Sertifikası, 2012, Samsun/TURKEY

Certificate of Attendance," Making Science Look Simple, 2011, Hessen/GERMANY

Certificate of Attendance," Making Science Look Simple, 2011, Mikulov/CZECH REPUBLIC

Certificate of Attendance," Making Science Look Simple", 2011, Buzau/ROMANIA

Certificate of Attendance, "Making Science Look Simple", 2010, Leicester/ENGLAND

Certificate of "Teaching and Learning for Democratic Citizenship Programme", 2008, Odense/DENMARK

Projects:
Ministry for EU Affairs, Comenius Project, "The Maths Rising All Europe" 2016-2018
Ministry for EU Affairs, Comenius Project, "Making Science Look Simple" 2010-2012

Contact:
E-mail: doganbaydal@gmail.com
Mobile Phone: +905544920832